ADVANCEMENT OF METHODS FOR PASSIVE ACOUSTIC MONITORING:

A Framework for the Study of Deep-Diving Cetacean

Ph.D Thesis

2018

Alba Solsona Berga

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Advancement of methods for passive acoustic monitoring: a framework for the study of deep-diving cetacean populations

Alba Solsona Berga

ADVANCEMENT OF METHODS FOR PASSIVE ACOUSTIC MONITORING:

A Framework for the Study of Deep-Diving Cetacean Populations

Alba Solsona Berga

Barcelona, Novembre 2018

Tesi presentada per a l'obtenció del títol de Doctora per a la Universitat Politècnica de Catalunya

Programa de Doctorat en Ciències del Mar

Departament d'Enginyeria Civil i Ambiental

Director de Tesi

Dr. Michel André

Laboratori d'Aplicacions Bioacústiques

Universitat Politècnica de Catalunya

Codirector de Tesi

Dra. Simone Baumann-Pickering

Scripps Institution of Oceanography

University of California San Diego

This work was partly conducted at the Laboratori d'Aplicacions Bioacústiques from January 2014 to June 2016 supported by the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR-SGR, 2014 – 2016), and completed at Scripps Institution of Oceanography from November 2016 to December 2018 supported by the Naval Facilities Engineering Command (NAVFAC) Cooperative Agreement grant, and National Oceanic and Atmospheric Administration (NOAA), Cooperative Institute for Marine Ecosystems and Climate (CIMEC), and Northeast Fisheries Scientific Center (NEFSC) Cooperative Agreement grant.

Graphic art: Soundscape main elements studied in this work. Original work by Judit Solsona Berga

Barcelona, November 2018

A tu padrina, per transmetre'ns la teva eterna curiositat

In Memoriam

Acknowledgments

This dissertation would not have been possible without a myriad of people to whom I show my sincere thanks for the academic and professional support and most importantly, for their affection. First of all, I take this opportunity to express my gratitude to my advisor, Simone Baumann-Pickering, for being there from the beginning and teaching me what it means to investigate throughout all these years. She has provided me with numerous amazing opportunities, and also with what I value the most, moral support not only for research but also for life. Mi más sincero agradecimiento a mi supervisor Michel André por darme la oportunidad de emprender este trayecto de tesis. No sólo he aprendido de ti un montón sobre los aspectos de liderazgo en proyectos científicos, sino sobre cómo hacer la ciencia algo más aplicada. I would like to thank John Hildebrand for giving me the incredible opportunity to come to Scripps and for believing in me. For those tireless afternoons wrestling with code and learning with enthusiasm from your vast knowledge. You have continued to encourage me to move forward and discover all the amazing things about bioacustics. I am also extremely grateful to Marie Roch for your patience and always taking time to teach me more about machine learning and best coding practices.

Gracias a toda la gente del LAB que en algún momento me ha tendido una mano: Marta Solé, Mike van der Schaar, Ludwig Houegnigan, Pablo Pla y Steffen de Vreese. Especialmente, gracias a Pablo y Steffen por estar allí dando apoyo en los momentos más confusos y por hacer los días en el LAB, un lugar familiar. Se echan de menos esas largas conversaciones sobre la vida y la ciencia al lado del mar y en la roca. Ludwig por siempre ofrecer tu ayuda y tus palabras para calmar las emociones.

I am fortunate to have been part of SIO and worked with the most talented and brightest people I could have wished for. Especially for the loudest cheerleaders in the world: Natalie Posdaljian, Anne Simonis, and Goldie Phillips, my words can hardly express my gratitude. Natalie, your help goes beyond the deep sea, sharing the office with you have been the best whale pod I could have ever ask for. Always ready to provide with the most fun snacks to endure the long working days, and fill them with smiles and fun conversations. I cannot wait to see what remarkable things you will discover about sperm whales in the coming years. Anne, thanks for being a sister, forever grateful for having you in my life. Thanks to the current and past grad students and staff in the lab, especially Jenny Trickey, Amanda Debich, Anna Krumpel, Rebecca Cohen, Regina Guazzo, Ashlyn Giddings, Eadoh Reshef, and Eric Snyder. To the infallible Dream Team (Jenny and

Katherine), for lighting up the deck with fun rituals and making sure that I did not miss any sperm whale encounter. I would also like to thanks Sean Wiggins, Erin O'Neill, Bruce Thayre, Beve Kennedy, and Sienna Thomas for the technical support, especially Bruce, who helped me even in his busiest times to make sure that all the computers were working smoothly. And special thanks are reserved for Kait Frasier, for being my partner in crime, I am glad we finally managed that the code made some sense.

Quiero dar las gracias, a Oliver, por apoyarme y animarme siempre a seguir mi rumbo, incluso cuando mi trabajo me llevara al otro lado del charco. Es un verdadero regalo haber conocido a mi segunda familia aquí en San Diego, gracias a todos, Nerea, Amalia, Carlos, Javi, Laura y Fabi, por acogerme en mi etapa de nómada, sé que no ha sido fácil aguantar mis mareas en estos últimos meses. Y, hablando ahora de neuronita, Nere, lo que siento no lo escribo aquí, que ya lo sabes bien tú. Pero solo decirte gracias por apoyarme siempre, por tú infinita paciencia y añadir salitre a todas las soleadas tardes. Roy, part of this is also thanks to you, for making sure I laughed my way to the end of this long journey.

No importa el lluny que pugui estar que sempre heu fet que la distància no existeixi, gràcies família i amics. Pares, Ju, sé que els interminables viatges amb la caravana a la muntanya són el motiu pel qual vaig començar aquesta aventura, i encara que segueixi viatjant lluny, no importa el què que el vostre gran afecte per motivar-me no té kilòmetres. Gràcies!

Abstract

Marine mammals face numerous anthropogenic threats, including fisheries interactions, ocean noise, ship strikes, and marine debris. Monitoring the negative impact on marine mammals through the assessment of population trends requires information about population size, spatiotemporal distribution, population structure, and animal behavior. Passive acoustic monitoring has become a viable method for gathering long-term data on highly mobile and notoriously cryptic marine mammals. However, passive acoustic monitoring still faces major challenges requiring further development of robust analysis tools, especially as it becomes increasingly used in applied conservation for long-term and large-scale studies of endangered or data deficient species such as sperm or beaked whales. Further challenges lie in the translation of animal presence into quantitative population density estimates since methods must control for variation in acoustic detectability of the target species, environmental factors, and for species-specific vocalization rates.

The main contribution of this thesis is the advancement of the framework for long-term quantitative monitoring of cetacean species, applied to deep-divers like sperm and beaked whales. Fully-automated methods were developed and implemented to different populations of beaked whales in different conditions. This provided insight into generalization capabilities of these automatic techniques and best practices. However, implementing these tool kits is not always practical, and alternative methods for additional data processing were developed to expeditiously serve multiple purposes including annotation of individual sounds, evaluation of data in order to provide a highly dynamic technique, and classification for quantitative monitoring studies. This work also presents the longest time series to date of sperm whale presence using passive acoustic monitoring for over seven years in the Gulf of Mexico. Echolocation clicks were detected and discriminated from other sounds to understand the spatiotemporal distribution and structure of the population. A series of steps were implemented to provide adequate parameters and characteristics of the target population for density estimation using an echolocation click-based method. This allowed for the study of the Gulf of Mexico's sperm whale population, providing significant progress towards the understanding of the population structure, distribution, and trends, in addition to potential long-term impacts of the well-known catastrophic *Deepwater Horizon* oil spill and other anthropogenic activities.

The emergence of innovative approaches for detecting the presence of marine mammals and documenting human interactions can provide insight into ecosystem change. These species can be used as sentinels of ocean health to ensure the conservation of their marine environment into the next epoch.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 4

List of Tables

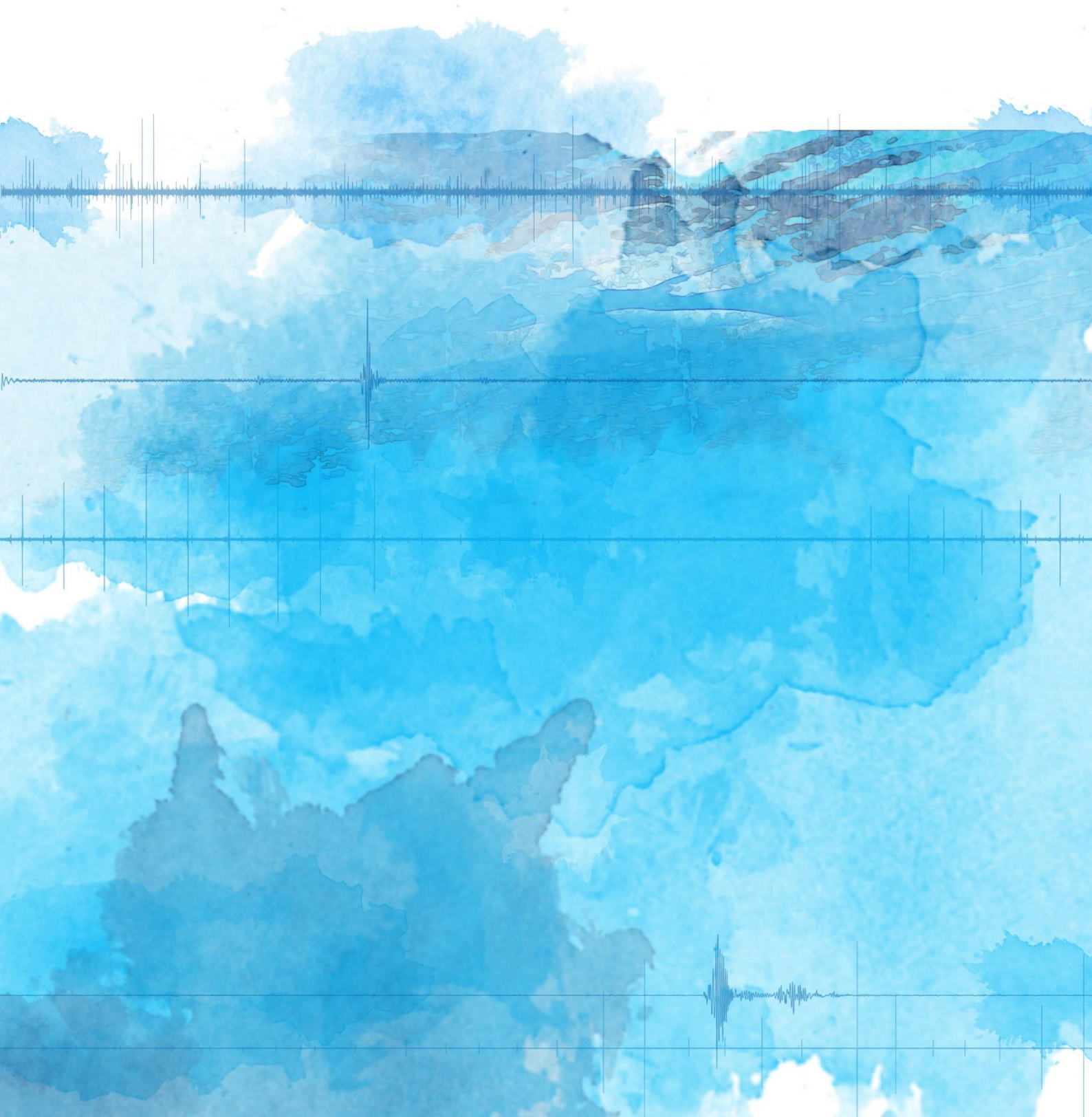
Chapter 1

Chapter 2

Chapter 4

List of Appendices

General Introduction



Innovative approaches for detecting odontocete presence

The concept of sentinel organisms suggests that understanding the status of some key organisms can provide an approach to evaluating ecosystem health. As top predators, marine mammals play a central role in maintaining and promoting healthy marine ecosystems and therefore, are strong candidates to serve as sentinel organisms. The influence of their ecological role on the oceans has been substantially undervalued (Roman et al., 2014). In recent years, sufficient understanding has emerged on their powerful and positive influence on ocean dynamics, global carbon pump and storage, and the health of commercial fisheries (Roman et al., 2014). There is evidence that monitoring the distribution, abundance, and density of these sentinel animals (Moore, 2008) can provide insight into short and long-term changes in ecosystems which guide current conservation efforts. Their pelagic and highly mobile nature has led to difficulties in studying their distributions and spatiotemporal behavioral patterns. Monitoring marine mammals by traditional ship or aerial surveys is time intensive, costly, and highly-dependent on external factors such as weather and daytime hours (Barlow et al., 2004). For deep-diving odontocetes, such as beaked and sperm whales, traditional survey methods are even less effective given the long periods of time spent beneath the surface foraging (Barlow and Gisiner, 2006; Tyack et al., 2006). Although ship based surveys are the most common marine mammal monitoring method, simply increasing the frequency of surveys may not result in more precise population assessment (Jewell et al., 2012).

Research on increasingly innovative approaches for studying marine mammals has led to alternative methods for detecting their presence and documentation of human impacts (Nowacek et al., 2016). Passive acoustic monitoring (PAM) is a feasible alternative to ship-based surveys for studying cetaceans who utilize the efficiency of sound under water to produce acoustic signals for communication with conspecifics, foraging, avoiding predators, and navigation. Deep-diving cetaceans are particularly reliant on sound when communicating, orienting, and hunting by echolocation as they regularly dive below 500 meters in near darkness (Schevill and McBride, 1956). Echolocation clicks can be detected, classified and used to track cetaceans (e.g., Watkins and Schevill, 1974; and Zimmer, 2011) through various PAM methods including autonomous gliders (Bingham et al., 2012; Klinck et al., 2012), towed hydrophone arrays (Barlow and Taylor, 2005; Norris et al., 1999), archival bottom-mounted acoustic instruments (Castellote et al., 2012; Mellinger et al., 2007; Wiggins and Hildebrand, 2007), and fixed cable observatories (André et al., 2011; Matsuo et al., 2013). The benefits of PAM include the ability to monitor mammals on large temporal and spatial scales, with extended instrument deployment periods, and the ability to operate in remote conditions regardless of weather or visibility (Mellinger et al., 2007). Limitations of PAM include that, in order to be detected, animals must be actively producing sound, and signals of interest have a limited detection range and must be isolated from other sounds (Zimmer, 2011). Regardless of these limitations, PAM technologies provide a non-

invasive method for collecting large volumes of data across time and space that can be applied to improve understanding of the mechanisms that drive cetacean spatial distributions and variability.

Manual detection and classification by trained acousticians are no longer viable with the recent, dramatic increase in the quantity of acoustic data that is being collected. Over the last several years, the focus of several research groups within the PAM community has been on the development of efficient methods for automatic detection and classification of recorded signals (Bittle and Duncan, 2013; Mellinger and Heimlich, 2013). Although several semi-automated algorithms have been utilized in long-term PAM studies (Baumann-Pickering et al., 2013; Hildebrand et al., 2015), fully-automated methods that incorporate algorithms which learn patterns and predict data, known as machine learning techniques, reduce the time needed to process long-term acoustic data and potential subjectivity (Frasier et al., 2017). These methods could expand population monitoring efforts of cetaceans by reducing the processing time needed to determine the conditions that drive animal presence, population dynamics, and animal behavior. The fully-automated classifiers' dependence on local area and animal population knowledge, limits their generalization capabilities, especially when applied to different locations or when conditions change over time. Given these restrictions, long-term monitoring using automatic classification techniques is not always practical and additional data processing techniques must be explored for efficient analysis of large datasets. Techniques prioritizing analysis speed that serve multiple purposes such as annotation of individual signals, evaluation of signal properties, and the ability to obtain false positive rates, serve as highly dynamic tools for expanding long-term monitoring capabilities.

Monitoring marine mammal populations with PAM technology has become more useful in the last decade with innovative statistical capabilities which convert the occurrence of vocalizations into a measure of population densities (Marques et al., 2009). Until the last decade, population assessment was estimated using methods based on visual observations which are disadvantageous for visually cryptic species or in adverse sighting conditions. The use of acoustic data to infer reliable estimation of population densities involves particular challenges (Mellinger et al., 2007). The relationship between detection events and actual animal densities is influenced by vocalization rate, variation in acoustic detectability, and environmental factors (Marques et al., 2013). First, the vocal behavior of a species must be well characterized before acoustics can be used for reliable density estimation. Second, detectability must be estimated using an adequate statistical method which, depending on the instrumentation and arrangement, will require a series of assumptions that can bias the estimation. Third, site and seasonal variability, including differences in environmental factors, must be accounted for. Most advances in acoustic density estimation have focused on cetacean species due to their reliance on sound for communication and navigation (Marques et al., 2013). However, the significant challenges listed above lead to

relatively few long-term acoustic monitoring of cetacean population densities (Frasier, 2015; Hildebrand et al., 2015). Therefore, we are witnessing an exciting time for opportunities to innovate acoustic technology capabilities, and development of methods to overcome the significant challenges associated with data management and analysis.

Thesis outline

The thesis emphases on the advancement of techniques related to acoustic data processing for long-term passive acoustic monitoring studies, provide insight into population structure and spatiotemporal distributions, and estimate population densities to infer population trends. This work represents a framework for long-term quantitative monitoring of cetacean species, applied to deep-divers like sperm and beaked whales. Ultimately, this research will contribute to open-source software for data analysis and facilitate future studies of sperm whale ecology.

- Chapter 1. Design and evaluation of an automated classifier for beaked whale echolocation clicks.

  A fully-automated method for processing acoustic data of beaked whale echolocation signals is described using machine learning techniques. It provides an evaluation of the performance on the generalization capabilities to process data from different regions and conditions. It describes best practices for addressing broader applicability of this methods.

- Chapter 2. *DetEdit*: A graphical user interface for annotating and editing events detected in acoustic data.

  A description of open-source software available for acoustic data analysis is provided which enables processing, visualizing, and annotating data. This makes data analysis highly dynamic and moves towards efficient and expeditious classification techniques for quantitative monitoring studies. It provides a description of the methodological technique developed to identify sperm whale echolocation signals used throughout the next chapters.

- Chapter 3. Population structure and patterns of habitat use of sperm whales in the Gulf of Mexico revealed by seven years of passive acoustic monitoring.

  A methodology is implemented to characterize population structure of sperm whale encounters through passive acoustic monitoring. This study is based on the comparison of acoustical patterns of echolocation clicks and measurements of acoustical animal length to provide possible maturity classes defining the population structure. It provides a description of spatiotemporal distributions and ultimately yields significant progress towards understanding sperm whale breeding migrations in the Gulf of Mexico.

- Chapter 4. Sperm whale density trends in the Gulf of Mexico over seven years of passive acoustic monitoring.

Adequate parameters and characteristics of the sperm whale population are determined based on findings from previous chapters and literature to provide density estimates using an echolocation click-based method. It provides a long-term time series of sperm whale densities to evaluate population trends and identifies the potential correlation between the population trends and the potential impacts from the *Deepwater Horizon* oil spill and other anthropogenic activities.

Publications of the results and contributions from others

All four chapters of the dissertation are currently being prepared for submission to scientific journals for publication. The dissertation author was the primary investigator and author. Co-author for these chapters will be Dr. Simone Baumann-Pickering who supervised the research.

The publication resulting from research outlined in chapter 1 will have Dr. Michel André, Dr. Mike van der Schaar, and Prof. Dr. Marie A. Roch as co-authors. They provided expertise in detection and classification algorithms, and valuable comments on the manuscript of this chapter.

Publications resulting from work presented in chapters 2, 3, and 4 will have Prof. Dr. John A. Hildebrand, Dr. Kait E. Frasier, and Natalie Posdaljian as co-authors. They provided expertise in signal detection and classification, data visualization, and density estimation methods, as well as valuable comments on the manuscript for these chapters.

References

André, M., van der Schaar, M., Zaugg, S., Houégnigan, L., Sánchez, a M., and Castell, J. V (**2011**). "Listening to the Deep: live monitoring of ocean noise and cetacean acoustic signals," Mar. Pollut. Bull., **63**, 18–26.

Barlow, J., and Gisiner, R. (**2006**). "Mitigating, monitoring and assessing the effects of anthropogenic sound on beaked whales," J. Cetacean Res. Manag., **7**, 239–249.

Barlow, J., Rankin, S., Zele, E., and Appler, J. (**2004**). "Marine mammal data collected during the Hawaiian islands cetacean and ecosystem assessment survey (HICEAS) conducted aboard the NOAA ships McArthur and David Starr Jordan, July - December 2002," NOAA Tech. Memo. NMFS,.

Barlow, J., and Taylor, B. L. (**2005**). "Estimates of Sperm Whale Abundance in the Northeastern Temperate Pacific From a Combined Acoustic and Visual Survey," Mar. Mammal Sci., **21**, 429–445.

Baumann-Pickering, S., Yack, T. M., Barlow, J., Wiggins, S. M., and Hildebrand, J. A. (**2013**). "Baird's beaked whale echolocation signals," J. Acoust. Soc. Am., **133**, 4321–31.

Bingham, B., Kraus, N., Howe, B., Freitag, L., Ball, K., Koski, P., and Gallimore, E. (**2012**). "Passive and active acoustics using an autonomous wave glider," J. F. Robot., **29**, 911–923.

Bittle, M., and Duncan, A. (**2013**). "A review of current marine mammal detection and classification algorithms for use in automated passive acoustic monitoring," Proc. Acoust. 2013,.

Castellote, M., Clark, C. W., and Lammers, M. O. (**2012**). "Acoustic and behavioural changes by fin whales (Balaenoptera physalus) in response to shipping and airgun noise," Biol. Conserv., **147**, 115–122.

Frasier, K. E. (**2015**). *Density estimation of delphinids using passive acoustics: A case study in the Gulf of Mexico.* Ph.D. Thesis, University of California San Diego, La Jolla, CA, 262 pages.

Frasier, K. E., Roch, M. A., Soldevilla, M. S., Wiggins, S. M., Garrison, L. P., and Hildebrand, J. A. (**2017**). "Automated classification of dolphin echolocation click types from the Gulf of Mexico," PLOS Comput. Biol., **13**, e1005823.

Hildebrand, J. A., Baumann-Pickering, S., Frasier, K. E., Trickey, J. S., Merkens, K. P., Wiggins, S. M., McDonald, M. A., et al. (**2015**). "Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico," Sci. Rep., **5**, 16343.

Jewell, R., Thomas, L., Harris, C., Kaschner, K., Wiff, R., Hammond, P., and Quick, N. (**2012**). "Global analysis of cetacean line-transect surveys: detecting trends in cetacean density," Mar. Ecol. Prog. Ser., **453**, 227–240.

Klinck, H., Mellinger, D. K., Klinck, K., Bogue, N. M., Luby, J. C., Jump, W. a., Shilling, G. B., et al. (**2012**). "Near-real-time acoustic monitoring of beaked whales and other cetaceans using a Seaglider," PLoS One, **7**, 1–8.

Marques, T. a., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. a., Moretti, D. J., Harris, D., et al. (**2013**). "Estimating animal population density using passive acoustics," Biol. Rev., **88**, 287–309.

Marques, T. a, Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. L. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales," J. Acoust. Soc. Am., **125**, 1982–94.

Matsuo, I., Akamatsu, T., Iwase, R., and Kawaguchi, K. (**2013**). "Automated acoustic detection of fin whale calls off Kushiro-Tokachi at the deep sea floor observatory," 2013 IEEE Int. Underw. Technol. Symp., IEEE, 1–3.

Mellinger, D. K., and Heimlich, S. L. (**2013**). "Introduction to the special issue on methods for marine mammal passive acoustics," J. Acoust. Soc. Am., **134**, 2381–2382.

Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., and Matsumoto, H. (**2007**). "An Overview of Fixed Passive Acoustic Observation Methods For Cetaceans," Oceanography, **20**, 36–45.

Moore, S. E. (**2008**). "Marine Mammals as Ecosystem Sentinels," J. Mammal., **89**, 534–540.

Norris, T. F., McDonald, M. A., and Barlow, J. (**1999**). "Acoustic detections of singing humpback whales (Megaptera novaeangliae) in the eastern North Pacific during their northbound migration," J. Acoust. Soc. Am., **106**, 506–514.

Nowacek, D. P., Christiansen, F., Bejder, L., Goldbogen, J. A., and Friedlaender, A. S. (**2016**). "Studying cetacean behaviour: new technological approaches and conservation applications," Anim. Behav., **120**, 235–244.

Roman, J., Estes, J. A., Morissette, L., Smith, C., Costa, D., McCarthy, J., Nation, J. B., et al. (**2014**). "Whales as marine ecosystem engineers," Front. Ecol. Environ., **12**, 377–385.

Schevill, W. E., and McBride, A. F. (**1956**). "Evidence for echolocation by cetaceans," Deep Sea Res., **3**, 153–154.

Tyack, P. L., Johnson, M., Soto, N. A., Sturlese, A., and Madsen, P. T. (**2006**). "Extreme diving of beaked whales," J. Exp. Biol., **209**, 4238–4253.

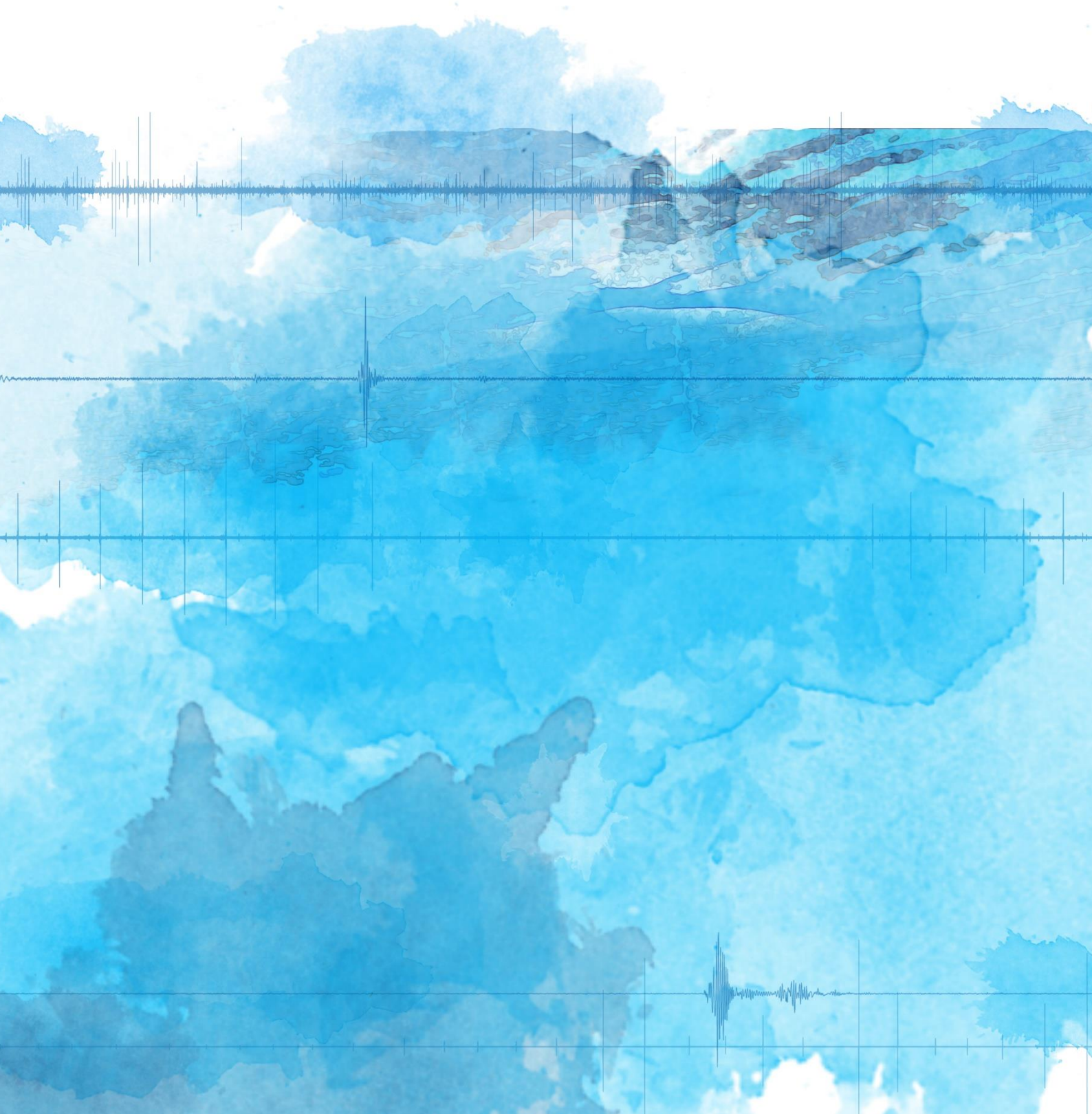
Watkins, W. A., and Schevill, W. E. (**1974**). "Listening to Hawaiian Spinner Porpoises, Stenella cf Longirostris, with a Three-Dimensional Hydrophone Array," J. Mammal., **55**, 319–328.

Wiggins, S. M., and Hildebrand, J. A. (**2007**). "High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring," Int. Symp. Underw. Technol. UT 2007 - Int. Work. Sci. Use Submar. Cables Relat. Technol. 2007, doi: 10.1109/UT.2007.370760.

Zimmer, W. M. X. (**2011**). *Passive acoustic monitoring of cetaceans*, Cambridge University Press.

Chapter 1

Design and evaluation of an automated classifier for beaked whale echolocation clicks



1.1. Abstract

The use of machine learning is increasing in the field of passive acoustic monitoring, and automated classifiers can improve our abilities to process the deluge of acoustic data. Knowing the limitations of these tools is essential for the effective use of algorithms to detect and classify signals produced by different species. Varying conditions (e.g. instrument sensitivity, ambient sound) can be found across long-term datasets or when using data from a number of locations. One challenge in the use of passive acoustics for species classification is the documentation of classifier performance under variable conditions. The performance of the classifier will largely depend upon its generalization capabilities. In this study, several features commonly used by human analysts were quantified, then incorporated into an automated process, and evaluated for their ability to discriminate beaked whale signals to species. An automated classification scheme based on Gaussian Mixture Models was trained using the manually identified click types provided by publicly available datasets from Detection, Classification, Localization and Density Estimation of Marine Mammals using Passive Acoustics workshops. The automated classifier was applied to datasets from the Pacific and Atlantic Ocean. A comparison framework to investigate classifier performance is provided by applying a bootstrap procedure on datasets from both regions. Both models showed similar classification performances with a slight improvement in the models from the same region, but performance between sites varied, particularly from the Atlantic data. These comparisons suggest that the application of classification methods should be developed knowing the area and nuances within populations that they can cover. It is also recommended to investigate potential population differences before developing a classifier to enhance their applicability.

1.2. Introduction

Monitoring species in their natural habitat is essential to determine the overall health and resilience of the marine ecosystem. As many species make extensive use of sound for a number of activities and vital behaviors, such as navigation, feeding and communication, passive acoustic monitoring (PAM) provides an alternative method to conventional visual surveys for population assessment and behavioral studies. Beaked whales are of particular interest as they remain one of the least understood groups of marine mammals, with relatively few abundance estimates in existence (Barlow et al., 2006; Baumann-Pickering et al., 2016; Hildebrand et al., 2015; Moore and Barlow, 2013). Anthropogenic underwater noise is now recognized as a global threat to marine fauna (Williams et al., 2015), and beaked whales appear especially vulnerable to high-intensity noise (Cox et al., 2006; Tyack et al., 2011) with numerous cases of mass-strandings following naval exercises in different locations (Arbelo et al., 2008; Balcomb III and Claridge, 2001; Filadelfo et al., 2009; Garrigue et al., 2016; Podestà et al., 2006). These elusive species are potential indicators for the effects of underwater noise, increasing the need for information about their spatiotemporal presence and abundance.

The beaked whale family currently includes 22 known species, constituting nearly one-fourth of cetacean species. They are challenging to study visually due to a deep-sea habitat, and extreme dive profiles with cryptic behavior at surface. For this reason, only three to four of the 22 species are reasonably well-known (Bianucci et al., 2008). Little was known about beaked whale species-specific vocalizations until recently. Research in the last decade provided acoustic descriptions for 9 of 22 species, Arnoux's beaked whale (*Beradius arnuxii*) (Rogers and Brown, 1999), Baird's (*Beradius bairdii*) (Baumann-Pickering et al., 2013c; Dawson et al., 1998), Northern bottlenose whale (*Hyperoodon ampulatus*) (Wahlberg et al., 2011), Longman's (*Indopacetus pacificus*) (Rankin et al., 2011), Blainville's (*Mesoplodon densirostris*) (Johnson et al., 2004, 2006; Madsen et al., 2005), Gervais' (*M. europaeus*) (Gillespie et al., 2009), Stejneger's (*M. stejnegeri*) (Baumann-Pickering et al., 2013b), Deraniyagala's (*M. hotaula*) (Baumann-Pickering et al., 2010), and Cuvier's (*Ziphius cavirostris*) (Zimmer et al., 2005) beaked whales. Other distinct beaked whale click types await species identification. Baumann-Pickering et al. (2014) suggested that the characteristic frequency modulated (FM) pulse types labeled as BW40, BW43, BW70, and BWC are produced by separate species, and overlap with sighting and stranding distribution of Hubb's (*M. carlhubbsi*), Perrin's (*M. perrini*), Pygmy (*M. peruvianus*), and ginkgo-toothed (*M. ginkgodens*) beaked whales, respectively, hypothesizing that these signals belong to these species. The signal type BWG was introduced and has strong similarity with BWC but is of unknown origin. Recently a different candidate signal BW37V was identified for Hubb's beaked whale by Griffiths et al. (2018) that has better evidence to link this signal to the species. Trickey et al. (2015) suggested that the BW29 FM pulse type could correspond to Southern bottlenose

whale (*H. planifrons*) and the BW37 FM pulse type to Gray's beaked whale (*M. grayi*). Other southern hemisphere click types have been described including BW58 (Baumann-Pickering et al., 2015), BW53 and BW39 (Giorli et al., 2018).

PAM has already demonstrated usefulness in documenting these visually elusive species. PAM is noninvasive and less subject to environmental conditions. However, the challenge lies in the processing and logging of the deluge of data produced from acoustic recorders. Continuous sound recordings with a sampling rate of 200 kHz, collected over one year, produces approximately 12 TB of data. Such large data from long-term passive acoustic recorders, often spanning multiple sites and multiple years, colloquially termed "big data," poses outstanding advantages to infer patterns of species presence, behavior, and density. Processing these data involves the generation of efficient methods for detecting and classifying echolocation transient signals. Machine learning techniques, such as *classification*, have the potential to classify echolocation clicks to species level or other criteria and reduce the analyst-guided processing time, to more quickly and reliably determine follow-up questions in regards to environmental, physical and biological drivers of animal presence, animal behaviors, or population dynamics.

There is an increase in use of acoustic detection and classification algorithms for beaked whale signals. Over the last years, semi-automated algorithms have been implemented for long-term PAM studies (Baumann-Pickering et al., 2013a, 2014; Hildebrand et al., 2015) and real-time monitoring using towed hydrophone arrays (Yack et al., 2013), seagliders (Klinck et al., 2012), and profiling floats (Matsumoto et al., 2013). However, these approaches require evaluation by experienced analysts to assist in signal discrimination which is time-consuming for processing long-term acoustic data and potentially subjective (Frasier et al., 2017). Efforts to classify small sets of odontocetes with fully automated algorithms have commonly included beaked whale species (Gillespie and Caillat, 2008; Klinck and Mellinger, 2011; Roch et al., 2008, 2011). These algorithms, together with three additional available beaked whale algorithms have been tested on one small dataset from towed array data that included several beaked whale species and one non-beaked whale species (Yack et al., 2010). The compared detectors were not trained on towed array data, albeit an overall acceptable performance was obtained. The method that combined cepstral feature vectors with Gaussian Mixture Models (GMM) (Roch et al., 2008, 2011) had the best correct detection rate despite having higher false detection rates than the three other algorithms. Recently, LeBien and Ioup, (2018) applied step-wise discriminant analysis to determine an optimal feature set for discriminating signals to species, and developed and evaluated several unsupervised classification methods for classification of beaked whale clicks in the Gulf of Mexico. The spectral clustering routine had the best results in the Gulf of Mexico data in comparison to other clustering algorithms (k-means, Ward hierarchical clustering and Chinese Whispers) and was tested with two publicly available recordings from the Bahamas, both 30

minutes in duration. Although representative clusters were identified from the encountered species, ground truth data was unknown.

Cuvier's beaked whale is one of the best-known species of the family Ziphiidae but are still poorly understood. Cuvier's have a cosmopolitan distribution in all oceans, with the exception of very high-latitude polar regions of both hemispheres (Jefferson et al., 1993; MacLeod et al., 2006). The Mediterranean population of Cuvier's beaked whale is genetically distinct between different oceanic basins and is subject to multiple conservation threats; the most significant factors being anthropogenic noise, fishery interactions, and shipping (Cañadas et al., 2018). Hence, there was a large incentive to generate an automated classifier suited to detect Cuvier's beaked whale vocalizations from the ANTARES (Astronomy with a Neutrino Telescope and Abyss environmental RESearch) platform in the Mediterranean Sea, which contains a system of acoustic hydrophones among other sensors. The audio stream from this cabled observatory is automatically processed in real-time by the "Listen to the Deep Ocean Environment" (LIDO, http://listentothedeep.com) software package. The LIDO software package consists of successive modules to automatically detect and classify impulsive sounds (André et al., 2011). It includes impulse and short tonal sound detectors applied to several frequency bands covering the full bandwidth of the data to detect impulsive shipping, sperm whales and delphinid vocalizations. Each detected impulse is systematically processed with a feature extraction step and assigned a class with a GMM based classifier (André et al., 2017; Houégnigan et al., 2010; Zaugg et al., 2010, 2012, 2013).

Here a set of relevant features to develop a classifier based on GMMs is presented and evaluated, which is complex enough to capture the particular characteristics of beaked whale species vocalizations. Previous manually identified recordings of the target species from the ANTARES platform were not available in order to develop the classifier. Instead, algorithms were trained and evaluated with two publicly available, large, annotated datasets. These datasets were also used to test the classifier's generalization capabilities and its potential for processing new data from different domains.

1.3. Methods

1.3.1. Datasets

Acoustic recordings were analyzed from several large datasets (total size: 12 TB) from different locations provided by the International Workshop on Detection, Classification, Localization and Density Estimation of Marine Mammals using Passive Acoustics (DCLDE); data from the Pacific for the DCDLE2015, and the Atlantic for the DCLDE2018. Multiple datasets were used to provide a better evaluation of the generality of the classification algorithm. The combined data covers a range of spatial, temporal, and recording variability, containing acoustic signals of 13

different odontocete species including four beaked whale species. The datasets were divided into development and validation sets. The development set was used to develop the classification models and the validation for evaluating the performance on novel data (**Table 1.1**).

1.3.2. Detection

The detection process in this study is based on the LIDO software that includes impulse detectors applied to several frequency bands covering the full bandwidth of data. While only a summary is given here, a detailed description is found in Zaugg et al. (2010). The audio data was processed in consecutive segments of 16.8 s length and band-pass filtered to attenuate energy outside of the desired frequency range. Here, the analysis focuses on two detection bands, 20 – 46 kHz and 35 – 80 kHz, which are relevant for the detection of beaked whale clicks. An estimate of the offset position of the click is given when the magnitude of the filtered signal, which is smoothed via running arithmetic means, exceeds a background noise threshold. The estimation of the background noise level was establish empirically (from previous data, Zaugg et al. (2010)) with a moving median taken over the same magnitude signal and multiplied by a factor of 2 to determine the threshold.

In order to obtain some of the features for the classification process, a second detection step was implemented to determine the exact start and end position of the click. From the estimated position, a 2 ms window was centered on the signal, and a 5^{th} order bandpass elliptic filter with a lower passband frequency of 5 kHz and a higher passband frequency of 80 kHz was applied. Subsequently, a finer resolution of the click position was estimated as presented by Soldevilla et al. (2008) and Roch et al. (2011) using the Teager energy operator (Kaiser, 1990; Kandia and Stylianou, 2006).

1.3.3. Feature Extraction

This section introduces the feature extraction process and selection of feature subsets for the classification of echolocation clicks. The main focus has been on developing and implementing reliable algorithms for feature extraction to discriminate beaked whale signals when processing real-time data from different ocean observatories. This requires algorithms to perform under a diversity of noise conditions.

1.3.3.1. Statistical Measures

Five mathematical equations were used to extract features from spectral and temporal energy contours of the clicks. The standard equations used in statistics were utilized to compute the center, dispersion (standard deviation), skewness, kurtosis, and Shannon entropy of random variables. They are used here in a purely descriptive way as in Zaugg et al. (2010).

Table 1.1. Summary data used for developing and evaluating classification models. Values indicate number of encounters per species and dataset (development data | validation data).

Site				Species												
Name	**Latitude**	**Longitude**	**Depth (m)**	**Zc**	**Bb**	**Mb**	**Me**	**Gg**	**Lo**	**La**	**Ssp**	**Pm**	**UDA**	**UDB**	**UD**	**UO**
Pacific dataset																
DCPP-C	35˚ 24.0' N	121˚ 33.8' W	1000	6|2	2|0			1|20	6|7			60|2				30|3
SOCAL-E	32˚ 39.4' N	119˚ 28.4' W	1300	108|63	2|1			2|0	3|3			8|5				31|7
SOCAL-R	33˚ 09.6' N	120˚ 00.6' W	1200	73|6	2|2			0|3	2|1			12|0				48|48
Subtotal				187|71	6|3			3|23	11|11			80|7				109|58
Atlantic dataset																
GOM-DT	25˚ 32.3' N	84˚ 37.9' W	1200	31|3			41|4	32|2			2|4				26|2	
HAT-A	35˚ 20.8' N	74˚ 50.9' W	840	22|5			25|0	1|0						2|0	84|36	
WAT-HZ	41˚ 03.7' N	66˚ 21.1' W	859	20|10		26|15		0|2		5|9			10|10	1|0	61|34	
WAT-NC	39˚ 49.9' N	69˚ 58.9' W	980	1|1		1|0	7|0	10|6		2|0			66|17	1|0	56|2	
Subtotal				74|19		27|15	73|4	43|10		7|9	2|4		76|27	4|0	227|74	
Total				161|90	6|3	27|15	73|4	46|33	11|11	7|9	2|4	80|7	76|27	4|0	227|74	109|58

Zc: *Ziphius cavirostris*; **Bb**: *Beradius bairdii*; **Mb**: *Mesoplodon bidens*; **Me**: *M. europaeus*; **Gg**: *Grampus griseus*; **Lo**: *Lagenorhynchus obliquidens*; **La***: L.acutus*; **Ssp**: *Stenella sp*.; **Pm**: *Physeter macrocephalus*; **UDA**: delphinid type A; **UDB**: delphinid type B; **UD**: unidentified delphinid; **UO**: unidentified odontocetes.

1.3.3.2. Individual Click Features

Several measurements were taken from individual clicks. Peak frequency, duration measured as the time interval between the click onset and click offset as presented by Roch et al. (2008) and Soldevilla et al. (2008). Duration of the envelope above upper-half calculated as the duration of the portion of the normalized envelope [0 >E < 1] above a threshold of E ≥ 0.5, and sweep rate taken as the frequency difference between start and end points of the click.

There is currently no standardized method for estimating the sweep rate of the characteristic frequency modulated beaked whale click. Some studies applied least-square fitting of the spectra to obtain the frequency difference (Baumann-Pickering et al., 2013a, 2013b, 2013c; Trickey et al., 2015) and others used linear and quadratic fit of the instantaneous frequency with a fix time window width (Baumann-Pickering et al., 2010; McDonald et al., 2009). The instantaneous frequency was chosen for this analysis. Obtaining the analytic signal by applying a Hilbert transform allows the differentiation of the phase and therewith calculate the instantaneous frequency (Huang et al., 1998). However, applied to bioacoustic data, it may result in negative instantaneous frequencies which are not useful in describing real physical phenomena like sound waves. Instead, the osculating circle method (OCM) developed by Hsu et al. (2011) was applied. OCM uses the Hilbert transform method with a modified definition of phase that ignores the location of the reference point. Occasionally, received clicks are overlaid with echoes or other clicks, causing distorted waveforms with discontinuities in the phase, also resulting in negative instantaneous frequencies. Following the study of Venkitaraman and Seelamantula, (2012) based on electromagnetic pulses, a bipolar threshold-pair in the instantaneous frequencies was used to locate the discontinuities and compensate for the jump size in the phase angle. After compensating for negative instantaneous frequencies, a linear fit was applied to obtain the frequency modulation gradient, estimated within a window width defined by a noise factor based on the sound pressure level of the segment and the signal-to-noise ratio. The goodness of fit was also taken as a classification feature.

1.3.3.3. Click Train Features

Odontocetes emit a sequence of several consecutive clicks called a click train. The distinctive rhythmic pattern of a click train, defined by a consistent inter-click-interval (ICI), has been one of the most relevant criteria for classification of echolocation signals to species level for beaked whales (Baumann-Pickering et al., 2013a) and other odontocetes (Castellote et al., 2013; Frasier et al., 2017; Roberts and Read, 2015). Although ICIs received at a sensor are subject to significant variability due to the orientation of the animal or its behavior in some cases, these changes modify received click properties and hence, a signal may not be associated with its sequence (Au, 2004; Gassmann et al., 2015; Simard et al., 2010). The rhythmic pattern remains nearly constant over

several consecutive clicks (Baumann-Pickering et al., 2013a) making it possible to automatically identify the sequences of clicks. Additionally, there are other processing challenges that increase the difficulty of this task, such as interleaved trains from several individuals or the presence of echoes from the clicks themselves, jittered time of arrival, and missing clicks not detected within the sequence. Previous methods computed ICIs by searching for a periodic pattern within similar clicks using one or several acoustic click descriptors. These methods range from simple correlation techniques (Bahl et al., 2002; Lepper et al., 2005; Starkhammar et al., 2011) to advanced methods such as artificial neural networks (Houser et al., 1999; Ioup et al., 2007), statistic clustering (Baggenstoss, 2011; Gervaise et al., 2010), and multi-hypothesis trackers (Gérard et al., 2008, 2009). Most of them are not adequate to quickly and reliably determine ICIs from long-term data and need relatively invariable parameters and/or a-priori training to match the rhythmic pattern. Rhythmic analysis algorithms have been applied successfully to improve the task of de-interleaving click trains and computing ICIs at a single sensor (Le Bot et al., 2015; Zaugg et al., 2013). The Zaugg et al. (2013) rhythmic analysis algorithm was applied, which uses a complex autocorrelation function by Nishiguchi and Kobayashi, (2000) that highlights the fundamental ICIs of the interleaved click trains to improve the subharmonics suppression resulting by standard autocorrelation functions (parameters in **Appendix IV,** detailed methodological description in Zaugg et al. (2013)). This method attaches information obtained from several nearby clicks with similar spectral properties to a focal click. In addition, three other classification features were obtained from the identified consistent sequences. A reliability measure – taken as an estimate of the number of clicks that could fit the computed ICI of the focal pulse within a defined tolerance region, the number of clicks analyzed to estimate the fit, and the sweep rate within the click train – calculated as the standard deviation using all sweep rates within the analyzed sequence.

1.3.3.4. Feature Subset Selection

In the interest of minimizing the computational cost of feature extraction necessary for a good classification performance, the performance of four different feature subsets combining some of the feature groups was inspected in the Pacific dataset (**Table 1.2**). Data exploration was used to find features that had better discrimination power (**Appendix IV**), and the features that explain lesser difference where excluded by feature groups. A final subset was created that encompassed all features regardless of their ability to individually discriminate between species.

1.3.4. Classification

Acoustic detections are not statistically independent because odontocetes produce many echolocation clicks during an encounter. To compensate, a classification methodology similar to that reported in Roch et al. (2011) was applied; where echolocation features were grouped by

Table 1.2. Overview of features and combination of feature subsets used for classification

	Subset					
Category	**A**	**B**	**C**	**D**	**E**	**Feature**
Spectrum statistical measures	x	x	x	x	x	Center
		x	x		x	Standard deviation
		x	x		x	Skewness
		x	x		x	Kurtosis
		x	x		x	Degree of peakiness (Shannon entropy)
Envelope statistical measures	x		x	x	x	Standard deviation
	x		x	x	x	Skewness
	x		x	x	x	Kurtosis
	x		x	x	x	Degree of peakiness (Shannon entropy)
Individual click	x	x	x	x	x	Duration
	x	x	x	x	x	Peak frequency
	x	x	x	x	x	Sweep rate
	x	x			x	The goodness of fit of sweep rate
	x	x	x	x	x	Duration envelope above upper-half
Click train	x	x	x	x	x	Inter-click-interval
	x	x			x	Reliability measure
	x	x	x	x	x	Standard deviation of sweep rate
	x	x			x	Number of clicks

encounters, defined as a set of clicks separated from another set by at least 5 min. All echolocation clicks with missing values in one of the extracted features were removed from analysis and values were standardized to a mean of zero and a unit variance to minimize the influence of large ranges between the feature values.

Each species was modeled with a GMM (using a full covariance matrix) in which the prediction is given as a likelihood score. The higher the score, the higher the confidence or certainty of the classification model that the encounter corresponds to the class. The likelihood scores for click encounters were determined as a joint likelihood of all clicks within the encounter, treated therefore as independent within a sequence. Consequently, yielding to an underestimation of the likelihood. The encounter was classified to the species associated with the model that had the maximum summed logarithmic likelihood score. Only single species encounters were used for this study to simplify the evaluation process.

Optimal GMM parameters for fitting the defined dataset were selected as follows. First, a single mixture component per GMM was trained with the training set obtained from the development data applying 30 Monte Carlo trials. This process was repeated, and the number of mixture components was varied until an arbitrary restricted maximum number of 30 components per

GMM was reached. Second, the best model conformation for each species model was chosen based on the Bayesian information criterion (BIC), which considers the likelihood of the fit and the number of mixture components (i.e., model complexity) to account for over-fitting.

1.3.4.1. Design of experiments

When training a classification model, overfitting is a common problem. This occurs when a model is constrained to represent the training data so tightly that it fails to accurately generalize to unseen data. To achieve proper classification of data from other ocean basins, it was prioritized to obtain a model with optimal generalization performance. Therefore, the development data was partitioned in a balanced manner into three-fold experiments as applied in Roch et al. (2011, 2015) to avoid building overfitted models. Feature groups – representing features from an encounter – were maintained together within the partitioning, and groups that did not form a complete set of 100 feature vectors were discarded. A bootstrap procedure was implemented to create the partitions and increase the variability of the training data by randomly sampling 85 % of the data without replacement. Two-folds were selected for training to fit the models and the remaining data as a test set to estimate prediction errors. These folds were shuffled, with one of the training folds moving to the role of test fold and vice versa until all three folds were tested. This process was repeated 100 times yielding 300 experimental trials (**Figure 1.1A**). The best 25 % of all models were selected based on error classification rates of the development dataset and tested on the validation dataset to assess the generalization performance (**Figure 1.1B**).

Furthermore, to evaluate how a given model built with a dataset from one region performed on a different domain, the models from the Pacific dataset were tested with the Atlantic dataset and vice versa. Same standardization of feature values in training folds was applied to the test fold. For highest precision and applicability, the classification predictions were quantified with respect to ground truth data that had been manually annotated by analysts. There are many criteria to quantify the performance of a classifier. The probability of misclassification, the *error rate*, is the most popular one. The error rate is calculated as the fraction of misclassified encounters with respect to the total number of encounters in the dataset.

Different performance metrics were also considered to evaluate if the classifier was falsely inferring the existence or absence of the beaked whale species. For this, the possible types of errors were differentiated by counting and compiling in a confusion matrix. When an encounter is correctly classified to its species class, it represents a *true positive* (TP), but when assigned to

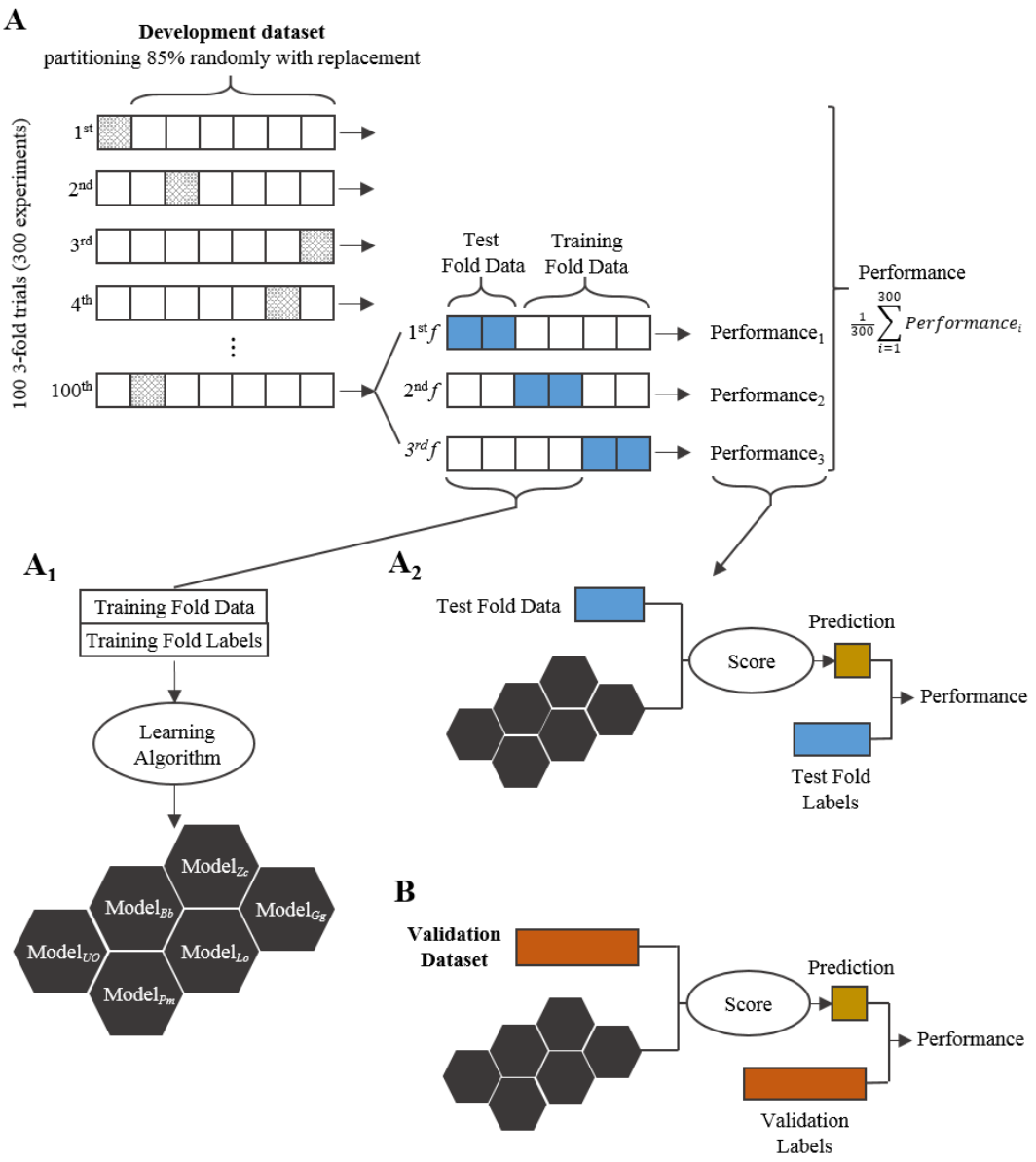


Figure 1.1. Workflow structure of the experimental setup for training, testing and evaluating classification models. Example of the Pacific dataset across 100 three-fold cross-validation trials.

a different class is a *false negative* (FN). If the encounter is from a different species class and is assigned to the target class, this represents a *false positive* (FP), and if assigned to a different class represents a *true negative* (TN). Several metrics can be derived from these definitions. *Precision* taken as the number of TP divided by the number of predicted positives, *Recall* or *sensitivity* as the number of TP divided by the number of actual positives, and *F-score* as

$$F = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

Here we used *precision*, *recall*, and *F-score* as a measure of a classifier's exactness, completeness, and balance between precision and recall, respectively. Another metric was defined to measure

percent time coverage of ground truth encounter data by the predicted encounter times, referred to as *truth coverage* and vice versa, the *detection coverage*.

1.4. Results

In this study, the performance of beaked whale species classification models was explored. The other classes (delphinid classes and sperm whale class) were used for multiclass classification purpose and are shown to indicate where the misclassification occurred. The target species was Cuvier's beaked whales (*Zc*), but results of two other beaked whale species, Sowerby's beaked whale (*Mb*) and Gervais' beaked whale (*Me*) from the Atlantic dataset are also shown. Baird's beaked whale (*Bb*) echolocation clicks were of a lower frequency than the other species (Baumann-Pickering et al., 2013c) and fall slightly outside of the frequency band of our detector and were consequently not further explored.

1.4.1. Development Dataset

1.4.1.1. Feature Set Selection

To select the best model conformation, five feature subsets of the Pacific development dataset were evaluated. Results of 300 experimental trials are shown in **Figure 1.2**. All classification rates for the Cuvier's beaked whale class with all different subsets of features were above 80 %. Subset A achieved the best performance with a 90.7 % classification rate, followed by subset D, E, C and B with classification rates of 89.4 %, 88.7 %, 85.5 %, and 81 %, respectively. Subset A was selected as the model configuration for further evaluations.

1.4.1.2. Classifier training and testing

Overall beaked whale classes (*Zc*, *Mb*, *Me*) in both datasets were recognized with above 80 % classification rates (**Figure 1.2**). A comparison of Subset A with class *Zc* on both the Pacific and Atlantic dataset showed a lower performance for the Atlantic data. Class *Zc*, when trained with Pacific and Atlantic data, had a 90.7 % and 83.4 % classification rate, respectively. Looking at misclassification rates for the *Zc* Pacific model, it was noticed that 6.7 % of Cuvier's beaked

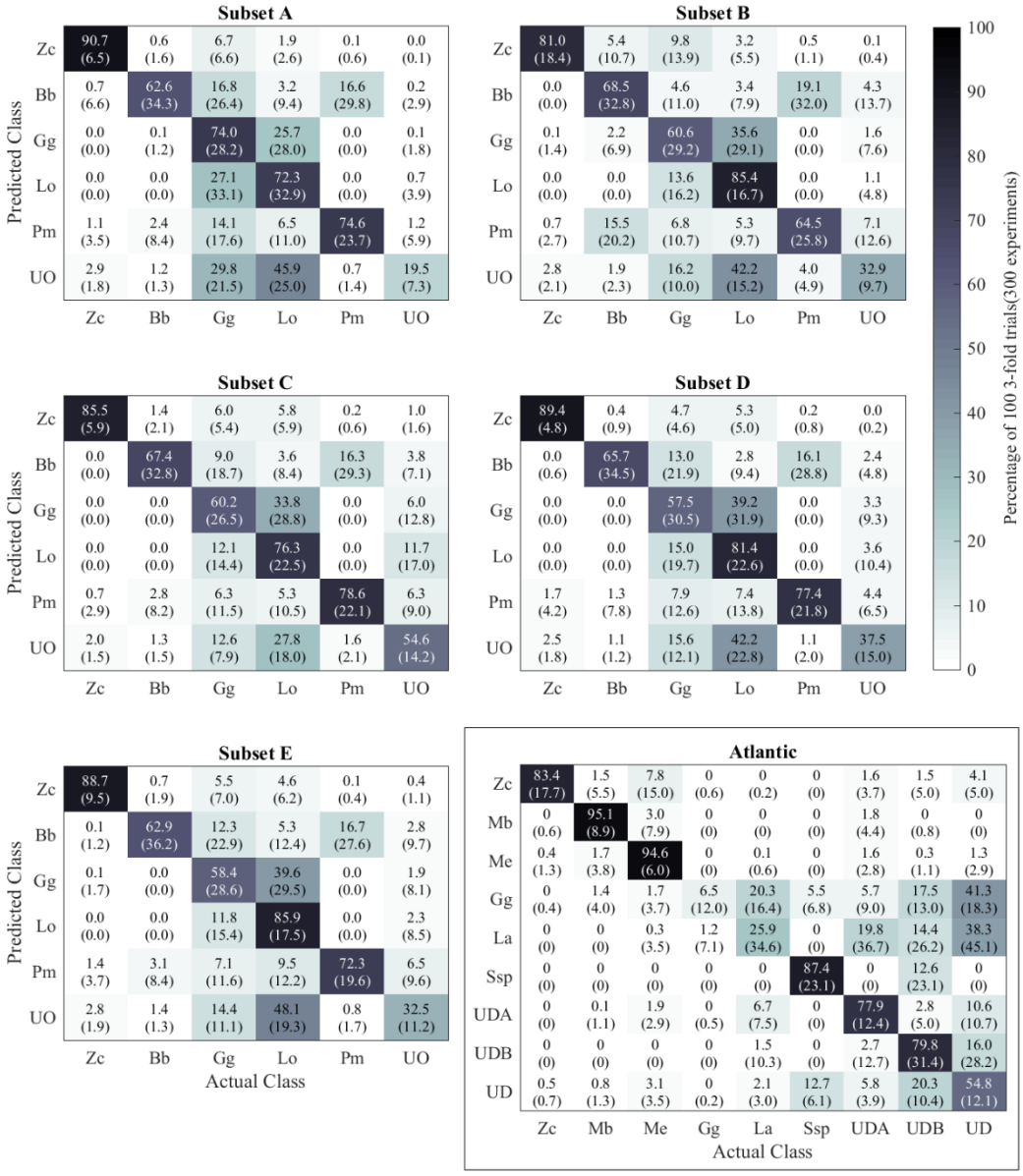


Figure 1.2. Mean confusion matrices (%) for the GMM classifier of 100 3-fold bootstrap experiments from the Pacific development dataset and from the Atlantic development dataset (boxed matrix). Each confusion matrix shows the classification performance using different subset of features, Atlantic confusion matrix is only using the features from Subset A. Elements of the matrix (numeric values and color-mapped visualization) represent the percentages of correctly (diagonal) or incorrectly (off-diagonal) classified species encounters by a GMM classifier during a 100 three-fold cross-validation procedure. Numbers in parenthesis are the standard deviation. Each column is normalized to a percentage, so values represent the percentages of 300 experiments for a given actual class that each possible class is predicted.

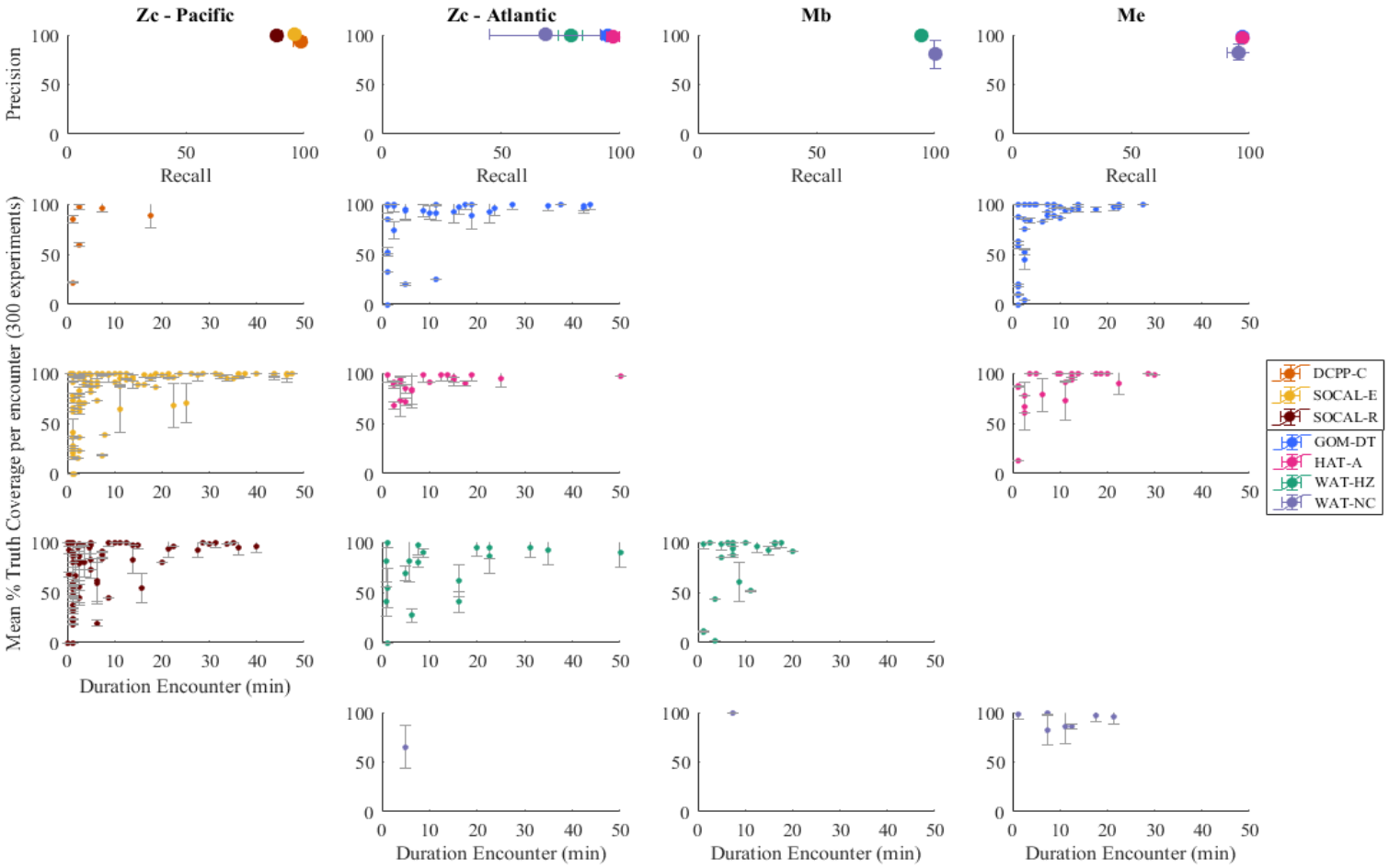


Figure 1.3. Evaluation measures of the classifiers' prediction performance for three species of beaked whales per site during a 100 three-fold Monte Carlo cross-validation procedure. Upper plots show the precision-recall (PR) averages and standard deviation per site. Bottom plot show, the percentage of ground truth coverage per site and encounter with respect to the encounter duration, mean, and standard deviation in each given encounter.

whale encounters among trials were identified as Risso's dolphin (*Gg*) and less than 2 % as Pacific white-sided dolphin (*Lo*). In contrast, the *Zc* Atlantic model did not misclassify any encounter as *Gg*, but it identified 7.8 % as *Me* and less than 5 % as unidentified dolphins (*UO*). Class *Mb* and *Me* are only present in the Atlantic dataset, so no comparison between datasets has been performed. However, excellent results were obtained for both classes in the Atlantic, class *Mb* was identified with a 95.1 % and *Me* with a 94.6 % classification rate.

Whether classification performance was accurate when compared to ground truth data was examined. Overall, the proportion of detected and correctly classified encounters by the models were similar for the three species, with precision above 90 % except for site WAT-NC, which had lower precision with 80.3 % (N_{enc} = 1) for class *Mb* and 82.7 % (N_{enc} = 7) for *Me* (**Figure 1.3**, **Appendix V**). Similarly, when analyzing the expected encounters detected by the correct model, a high performance for the three classes was reached, with a recall above 90 %; except for class *Zc* for sites SOCAL-R, WAT-HZ, and WAT-NC with 88.5 % (N_{enc} = 73), 79.4% (N_{enc} = 20), and 68.7 % (N_{enc} = 1), respectively (**Figure 1.3**, **Appendix V**). Together, these findings indicate that features from subset A are capable of performing a high classification precision for the three species of beaked whales found in the datasets for most sites. The precision for site WAT-NC was the most unsteady; however, the number of encounters available for this site was very low for all three species so conceivably models had a larger variability for this site.

In addition, the matching times between detected encounters and ground truth data were compared. Coverage was high for the long encounter (> 10min) at all sites with more than an 80 % of encounter time matching (**Figure 1.3**, **Appendix V**). On the other hand, for short encounters coverage decreased and varied between sites with a range of 4 – 16 % of the encounters covered less than 50 % by the classifier. The site WAT-HZ had the lowest coverage, with five encounters (N_{enc} = 26) and a match below 60 % between the classifier and the ground truth data.

1.4.2. Validation Dataset

To determine whether Cuvier's beaked whale classification models generalize for unseen acoustic data, both models obtained from the Pacific and Atlantic data were compared with the validation datasets of both ocean regions. Only the best 25 % of models among the 300 experimental trials were used to evaluate the validation data, resulting in 75 experimental trials. Overall, for the three Pacific sites, a good performance was obtained from both models, with a high precision-recall and F-scores above 80 % (**Figure 1.4**, **Appendix VI**). The Atlantic model produced slightly higher precision while the Pacific model produced higher recall rates, except for site DCPP-C where both models had the same performance. SOCAL-E is the site that had the lowest precision of nearly 80 % for both models. Similar misclassifications were obtained with 50 % of Baird's beaked whale (*Bb*) encounters and 30 % of the sperm whale (*Pm*) encounters classified as

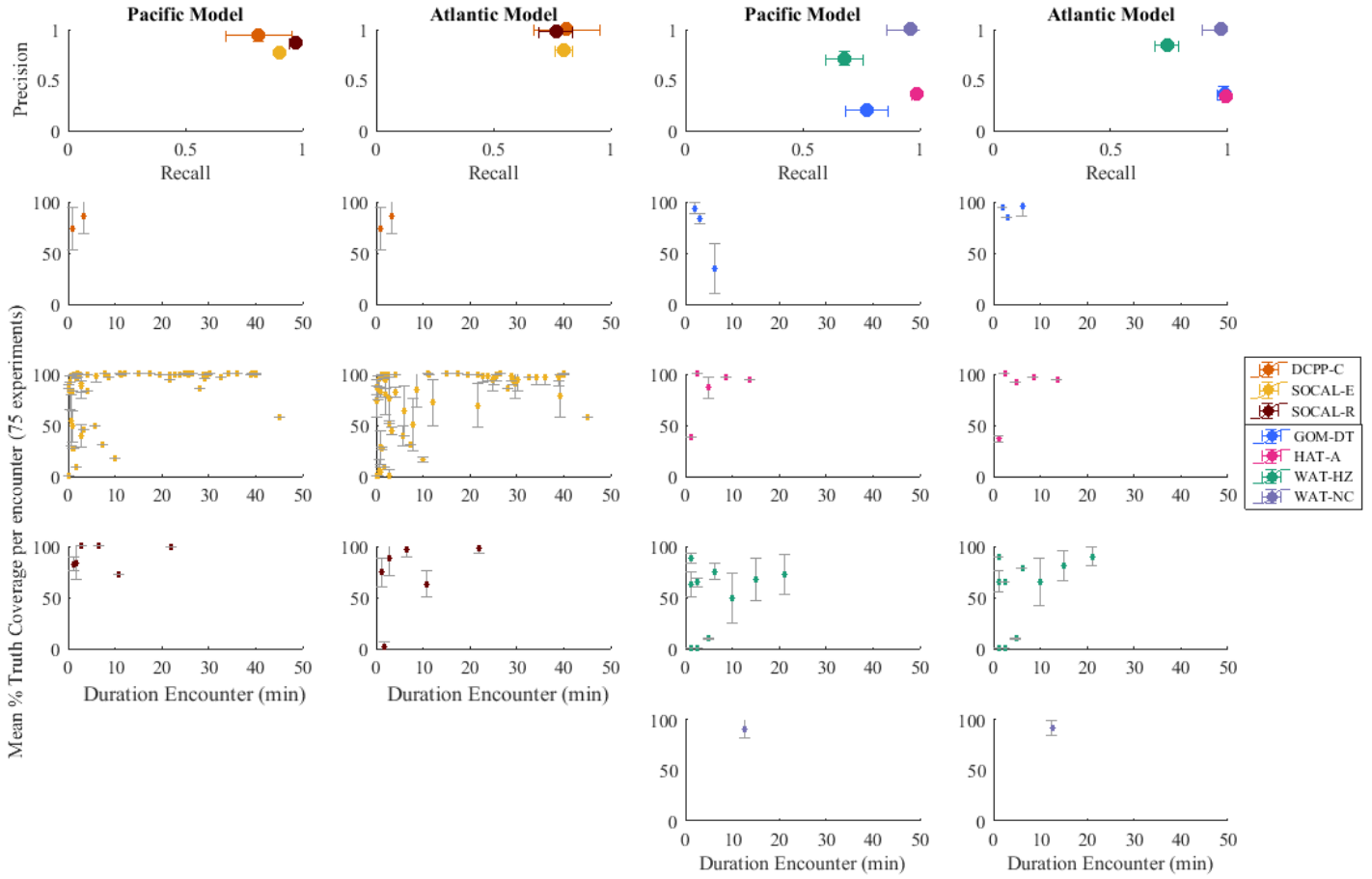


Figure 1.4. Evaluation measures of the classifiers' prediction performance for Cuvier's beaked whales per dataset for 75 experimental trials. Upper plots show the precision-recall (PR) averages and standard deviation per site. The bottom plots show, the percentages of ground truth coverage per site and encounter with respect to the encounter duration, mean, and standard deviation in each given encounter.

Cuvier's beaked whales (**Appendix II**). Mean truth coverage was high for most encounters, except for SOCAL-E with the Pacific model that resulted in a match lower than 60 % for 18 % of the encounters (N_{enc} = 63). In addition, 2 of 30 long encounters (above 10 min) had a low coverage, while the Atlantic model coverage decreased 8.5 % compared to the Pacific model. For SOCAL-R only one short encounter out of 6 had a low match when applying the Atlantic model. With the Atlantic data, there were significant differences between sites, with a precision below 40 % for both models at site GOM-DT and HAT-A and between 70 – 100 % for WAT-HZ and WAT-NC. Confusion matrices (**Appendix II** -**Appendix III**) do not show any relevant misclassification that could reduce the precision of both models. Therefore, we looked at the detection coverage instead of the truth coverage, and found that for these two sites, the match of the ground truth data with the predicted detections was below 20 % indicating that there was a high percentage of unlabeled encounters being detected by the detector. Mean percentage detection coverage and standard deviation for the Pacific model for site GOM-DT, HAT-A, WAT-HZ, and WAT-NC were 13.5 ± 26.8σ, 20.2 ± 36.1σ, 66.3± 45.2σ and 100.0 ± 0.0σ. Percent detection coverage for the Atlantic model were 21.2 ± 30.4σ, 19.3 ± 35.5σ, 81.6 ± 35.5σ, and 100.0 ± 0.0σ. In comparison, the recall was very high for most of the sites with a rate above 90 %, except for site WAT-HZ that ranged between 68 – 74 % for both models. For both cases, three encounters (N_{enc} = 10) had a low match below 20 %. The Pacific model for site GOM-DT had a lower recall with 21.4 % less than the Atlantic model, explained by one of the three encounters that had a low coverage. Together, these results suggest that both models are capable of classifying Cuvier's beaked whales when trained with data from vocalizations of another population or region, with no substantial differences of performance observed between both models. However, variability was observed between sites within the same region. These analyses may not be representative of the overall performance due to a low sample size of encounters per species for the validation dataset; therefore, models with the development dataset from the other domain were tested as an independent set to evaluate the performance with a larger sample size and foresee what the predictability of the models could be if applied to the ANTARES station. Both precision and recall varied significantly between sites and datasets (**Figure 1.5**, **Appendix VII**). For the Pacific sites, 20 % or below of Cuvier's beaked whale encounters at DCPP-C (N_{enc} = 6), and SOCAL-R (N_{enc} = 73) sites were classified as *Me*. This misclassification also occurred with a lower rate for the Atlantic models with the Atlantic development dataset (**Appendix III**). One particular misclassification case of *Lo* encounters classified as *Zc* were observed for SOCAL-R site on development data which lowered the recall. However, this was not the case for any of the other sites and test.

For Atlantic sites, a similar performance as seen in the validation test was observed for GOM-DT and HAT-A, with a low precision of ~30 % (**Figure 1.5**). However, the classification was notably different between development and validation data as shown in **Appendix II**. The Pacific *Zc* model

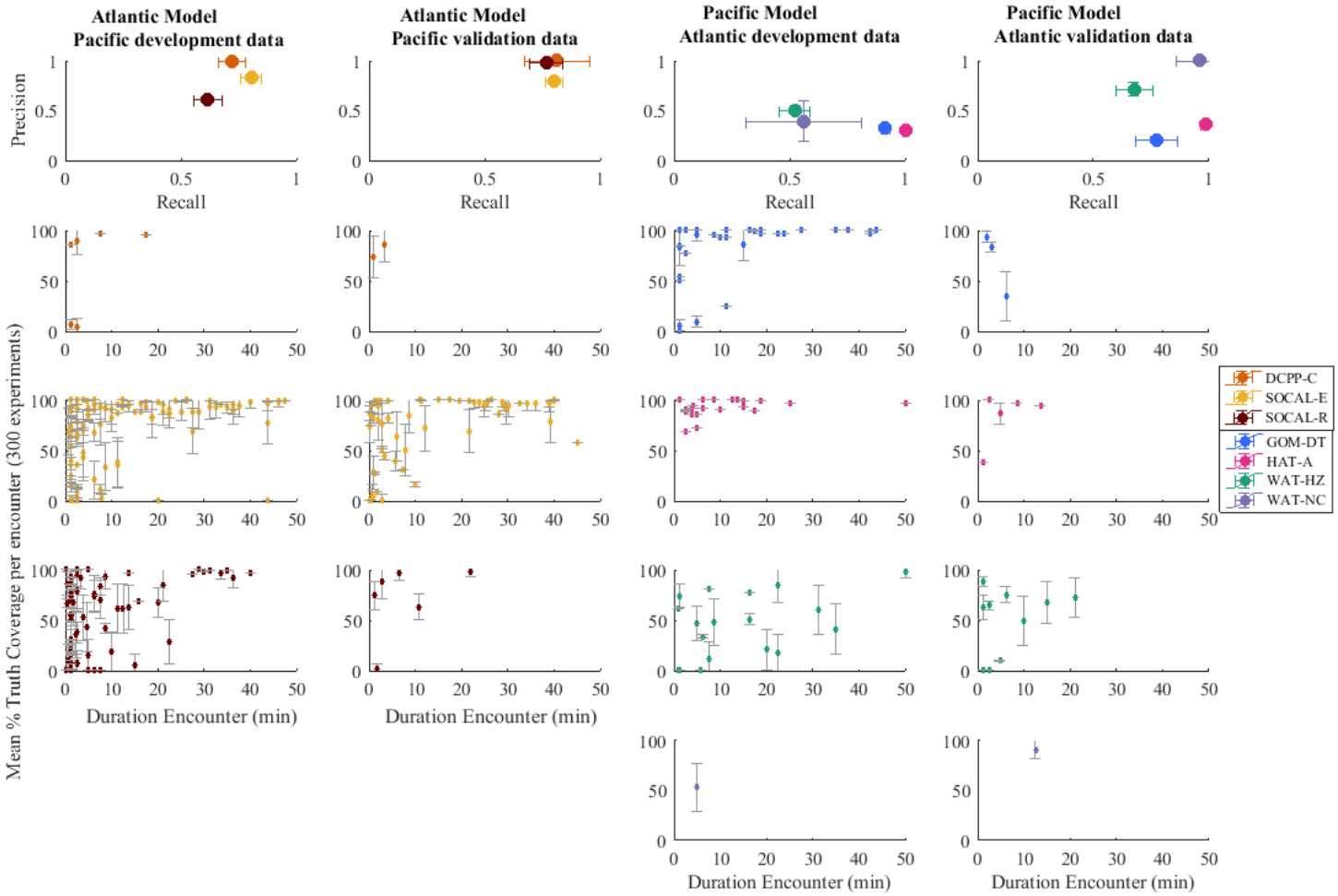


Figure 1.5. Evaluation measures of the classifiers' prediction performance for Cuvier's beaked whales per dataset for 75 experimental trials. Upper plots show the precision-recall (PR) averages and standard deviation per site. The bottom plots show, the percentages of ground truth coverage per site and encounter with respect to the encounter duration, mean, and standard deviation in each given encounter.

had a low false positive rate on validation data, whereas on development data there were more false positives. In this case, the development data had a better performance from a multiclass classification perspective. It was already assumed that the *Bb* model would not classify well on other beaked whale species because the frequency range is lower than the other species. Keeping in mind that the Pacific models only had the *Zc* and *Bb* models for beaked whale species, this leaves one model class to identify beaked whale data. Therefore, false positives for beaked whale species encounters from the *Zc* model were here evaluated as a good performance. On the contrary, recall lowered significantly to a rate of 30 – 44 % for WAT sites due to misclassification of Cuvier's encounters as *Gg* (**Appendix II**) and truth coverage showed the same trend as the validation tests. Finally, we tried the models on the ANTARES station, but no detections of Cuvier's beaked whale vocalizations were obtained.

1.5. Discussion

Automated analysis tools for processing acoustic data have become a necessity with the dramatic increase in quantity of data being collected by passive acoustic monitoring. These tool kits evolved quickly due to innovations in signal processing and machine learning (Digby et al., 2013; Klay et al., 2010; Roch et al., 2016). Applicability to monitor species populations with an automatic framework faces several challenges that can limit the efficacy of these methods for methodological monitoring purposes, and therefore affect the generalization capabilities which were evaluated in this study using an automated classification based on GMM models.

The classification process involves extracting features from the vocalizations of the target species typically describing its spectral and temporal characteristics (Baumann-Pickering et al., 2013a; LeBien and Ioup, 2018). It was found in this study that feature subset A containing temporal, click train, and specific spectral features, were capable of discriminating beaked whale sounds accurately to species, supporting the idea that these typical features combined can achieve a remarkably reliable performance to discriminate beaked whale species. Most importantly, that these features can be extracted and implemented in a fully automatic process. The majority of studies by other authors have not questioned how the discrimination of beaked whales with fully automated algorithms are capable of generalizing classifications with different recordings and conditions (Gillespie and Caillat, 2008; Klinck and Mellinger, 2011; LeBien and Ioup, 2018; Roch et al., 2011). A bootstrap procedure was implemented to enhance the generalization abilities of the GMM models when applied to several test problems. Two different models from different domains were trained and showed how these models performed on unseen data while comparing their results to the same sets. Evaluation of the development data pointed towards high discrimination of the three species of beaked whales. All tests had good coverage on encounters of longer duration (>10min long), albeit a trivial number of short encounters had unreliable

coverage. These results highlight the need for models to be trained with short, faint, and less distinctive encounters to improve the classification performance.

The models developed on the Pacific and Atlantic data performed well for most sites on the validation data. The exceptions were sites GOM-DT and HAT-A, where a significant number of new encounters were detected but not annotated, reducing the precision rate. Although better results were obtained with models from the same context, results indicate that both models have high potential to classify Cuvier's beaked whales when trained with data from different domains. Finally, the models were applied to the development data from the other domain to evaluate performance with more encounters and therefore, provide better interpretability of the generalization capabilities. Interpretation of the results was ambiguous due to the effects of the multiclass classification. Multiclass classification requires samples from both target and non-target classes to determine respective decision boundaries. If class models are different from the classes on the dataset, it can produce misleading results that need to be evaluated with caution. As an alternative to reducing the complexity of the classification, a threshold per class could be defined. Choosing an appropriate threshold, however, is not an easy problem in practice.

The models had an acceptable performance when classifying Cuvier's beaked whales even when trained with data from vocalizations of another region. No substantial differences in performance between either model were observed, but between sites within a region. Unfortunately, no Cuvier's beaked whale were detected in ANTARES using these models. Recently, Cañadas et al. (2018) provided predicted values of Cuvier's beaked whales for the region around ANTARES station using habitat modeling. They showed a very low probability of detection in this area, which could explain the lack of detection from the models in this station.

1.5.1. Recommendations

It is conceivable that model performance could change if applied to different locations or long-term time series. The ability to transfer knowledge to new conditions is a state-of-the-art concept in machine learning referred to as transfer learning (Torrey and Shavlik, 2009). The application of these methods could potentially improve the ability to generalize to conditions that are different from the ones encountered during training. Site recording conditions (Roch et al., 2015) of long-term monitoring studies and acoustic population characteristics will limit the generalization capabilities of the classifiers. The class discrimination power of some features could be heavily dependent on the context. Models will need to be developed knowing the area and population that they apply to. Other tools like unsupervised learning could be advantageous investigating possible differences among recordings and populations, and acoustic species population differences should be investigated.

1.6. Acknowledgments

I thank the numerous archival data contributors to this project and their respective funding agencies. Atlantic data was available with support through Sophie Van Parijs, NOAA-NEFSC, funding from NOAA and BOEM, contributor Danielle M. Cholewiak; Andrew J. Read, Duke University, funding from U.S. Fleet Forces Command under the U.S. Navy's Marine Species Monitoring Program, project management by Naval Facilities Engineering Command Atlantic, Joel T. Bell, contributors Lynne E.W. Hodge, Joy E. Stanistreet. Collection and analysis of the Pacific data was supported by the US Office of Naval Research, Michael J. Weise; US Navy Living Marine Resources Program, CNO-N45, Frank Stone, Bob Gisiner, Anurag Kumar; Naval Postgraduate School, C. Collins, J. Joseph, US Pacific Fleet, Chip Johnson. Gulf of Mexico data with support through Melissa Soldevilla, (Lance P. Garrison), NOAA-SEFSC, funding from NOAA-SEFSC, Steven Murawski through the Gulf of Mexico Research Initiative's CIMAGE consortium, the Natural Resource Damage Assessment partners, the US Marine Mammal Commission, and Michael J. Weise with the Office of Naval Research. I thank members of the SIO Whale Acoustics Laboratory, Jennifer Trickey, Kaitlin E. Frasier, Bruce Thayre, Erin O'Neill and John Hildebrand for data processing and providing species annotations. Computing resources for this work was available through Mike Weise, Office Naval Research.

1.7. References

André, M., Caballé, A., van der Schaar, M., Solsona, A., Houégnigan, L., Zaugg, S., Sánchez, A. M., et al. (**2017**). "Sperm whale long-range echolocation sounds revealed by ANTARES, a deep-sea neutrino telescope," Sci. Rep., **7**, 45517. doi:10.1038/srep45517

André, M., van der Schaar, M., Zaugg, S., Houégnigan, L., Sánchez, a M., and Castell, J. V (**2011**). "Listening to the Deep: live monitoring of ocean noise and cetacean acoustic signals," Mar. Pollut. Bull., **63**, 18–26. doi:10.1016/j.marpolbul.2011.04.038

Arbelo, M., De Quiros, Y. B., Sierra, E., Méndez, M., Godinho, A., Ramírez, G., Caballero, M. J., et al. (**2008**). "Atypical beaked whale mass stranding in almeria's coasts: Pathological study," Bioacoustics, **17**, 294–297. doi:10.1080/09524622.2008.9753853

Au, W. W. L. (**2004**). "The Dolphin Sonar: Excellent Capabilities In Spite of Some Mediocre Properties," AIP Conf. Proc., **728**, 247–259. doi:10.1063/1.1843019

Baggenstoss, P. M. (**2011**). "Separation of sperm whale click trains for multipath rejection," J. Acoust. Soc. Am., **129**, 3598–3609. doi:10.1121/1.3578454

Bahl, R., Ura, T., and Fukuchi, T. (**2002**). "Towards Identification of Sperm Whales from Their Vocalizations," SEISAN KENKYU, **54**, 409–13. doi:10.11188/seisankenkyu.54.409

Balcomb III, K. C., and Claridge, D. E. (**2001**). "A mass stranding of cetaceans caused by naval sonar in the Bahamas," Bahamas J. Sci.,.

Barlow, J., Ferguson, M., Perrin, W., Gerrodette, T., Joyce, G., Macleod, C., Mullin, K., et al. (**2006**). "Abundance and densities of beaked and bottlenose whales (family Ziphiidae)," J. Cetacean Res. Manag., **7**, 263–270. doi:10.1111/j.1748-7692.1997.tb00656.x

Baumann-Pickering, S., McDonald, M. A., Simonis, A. E., Solsona Berga, A., Merkens, K. P. B., Oleson, E. M., Roch, M. A., et al. (**2013a**). "Species-specific beaked whale echolocation signals," J. Acoust. Soc. Am., **134**, 2293–301. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23967959

Baumann-Pickering, S., Roch, M. A., Brownell, R. L., Simonis, A. E., McDonald, M. A., Solsona Berga, A., Oleson, E. M., et al. (**2014**). "Spatio-temporal patterns of beaked whale echolocation signals in the North Pacific," PLoS One, **9**, e86072. doi:10.1371/journal.pone.0086072

Baumann-Pickering, S., Simonis, A. E., Wiggins, S. M., Brownell, R. L., and Hildebrand, J. a. (**2013b**). "Aleutian Islands beaked whale echolocation signals," Mar. Mammal Sci., **29**, 221–227. doi:10.1111/j.1748-7692.2011.00550.x

Baumann-Pickering, S., Trickey, J. S., Wiggins, S. M., and Oleson, E. M. (**2016**). "Odontocete

occurrence in relation to changes in oceanography at a remote equatorial Pacific seamount," Mar. Mammal Sci., , doi: 10.1111/mms.12299. doi:10.1111/mms.12299

Baumann-Pickering, S., Wiggins, S. M., Roth, E. H., Roch, M. A., Schnitzler, H.-U., and Hildebrand, J. A. (**2010**). "Echolocation signals of a beaked whale at Palmyra Atoll," J. Acoust. Soc. Am., **127**, 3790–9. doi:10.1121/1.3409478

Baumann-Pickering, S., Yack, T. M., Barlow, J., Wiggins, S. M., and Hildebrand, J. A. (**2013c**). "Baird's beaked whale echolocation signals," J. Acoust. Soc. Am., **133**, 4321–31. doi:10.1121/1.4804316

Baumann-PickeringS., Širović, A., Trickey, J. S., Hildebrand, J. A., Reyes, M. V. R., Melcon, M. L., and Iñiguez., M. A. (**2015**). "Cetacean presence near Elephant Island, Antarctica, based on passive acoustic monitoring," Int. Whal. Comm. Sci. Comm. Doc. SC/66a/SH/18., San Diego, CA, USA.

Bianucci, G., Post, K., and Lambert, O. (**2008**). "Beaked whale mysteries revealed by seafloor fossils trawled off South Africa," S. Afr. J. Sci., **104**, 140–142.

Le Bot, O., Mars, J. I., Gervaise, C., and Simard, Y. (**2015**). "Rhythmic analysis for click train detection and source separation with examples on beluga whales," Appl. Acoust., **95**, 37–49. doi:10.1016/j.apacoust.2015.02.005

Cañadas, A., Aguilar de Soto, N., Aissi, M., Arcangeli, A., Azzolin, M., B-Nagy, A., Bearzi, G., et al. (**2018**). "The challenge of habitat modelling for threatened low density species using heterogeneous data: The case of Cuvier's beaked whales in the Mediterranean," Ecol. Indic., **85**, 128–136. doi:10.1016/j.ecolind.2017.10.021

Castellote, M., Leeney, R. H., O'Corry-Crowe, G., Lauhakangas, R., Kovacs, K. M., Lucey, W., Krasnova, V., et al. (**2013**). "Monitoring white whales (Delphinapterus leucas) with echolocation loggers," Polar Biol., **36**, 493–509. doi:10.1007/s00300-012-1276-2

Cox, T. M., Ragen, T. J., Read, a J., W, B. R., Balcomb, K., Barlow, J., Caldwell, J., et al. (**2006**). "Understainding the impact of anthropogenic sound on beaked whales," J. Cetacean Res. Manag.,.

Dawson, S., Barlow, J., and Ljungblad, D. (**1998**). "Sounds recorded from Baird's beaked whale, Berardius bairdii," Mar. Mammal Sci., **14**, 335–344.

Digby, A., Towsey, M., Bell, B. D., and Teal, P. D. (**2013**). "A practical comparison of manual and autonomous methods for acoustic monitoring," Methods Ecol. Evol., , doi: 10.1111/2041-210X.12060. doi:10.1111/2041-210X.12060

Filadelfo, R., Mintz, J., Michlovich, E., D'Amico, A., Tyack, P. L., and Ketten, D. R. (**2009**).

"Correlating military sonar use with beaked whale mass strandings: What do the historical data show?," Aquat. Mamm., **35**, 435–444. doi:10.1578/AM.35.4.2009.435

Frasier, K. E., Roch, M. A., Soldevilla, M. S., Wiggins, S. M., Garrison, L. P., and Hildebrand, J. A. (**2017**). "Automated classification of dolphin echolocation click types from the Gulf of Mexico," PLOS Comput. Biol., **13**, e1005823. doi:10.1371/journal.pcbi.1005823

Garrigue, C., Oremus, M., Dodémont, R., Bustamante, P., Kwiatek, O., Libeau, G., Lockyer, C., et al. (**2016**). "A mass stranding of seven Longman's beaked whales (Indopacetus pacificus) in New Caledonia, South Pacific," Mar. Mammal Sci., **32**, 884–910. doi:10.1111/mms.12304

Gassmann, M., Wiggins, S. M., and Hildebrand, J. A. (**2015**). "Three-dimensional tracking of Cuvier's beaked whales' echolocation sounds using nested hydrophone arrays," J. Acoust. Soc. Am., **138**, 2483–2494. doi:10.1121/1.4927417

Gérard, O., Carthel, C., and Coraluppi, S. (**2009**). "Classification of odontocete buzz clicks using a multi-hypothesis tracker.,"

Gérard, O., Coraluppi, S., Carthel, C., Undersea, N., Bartolomeo, V. S., and Spezia, L. (**2008**). "Analysis and Classification of Beaked Whale Buzz Clicks.,"

Gervaise, C., Barazzutti, a., Busson, S., Simard, Y., and Roy, N. (**2010**). "Automatic detection of bioacoustics impulses based on kurtosis under weak signal to noise ratio," Appl. Acoust., **71**, 1020–1026. doi:10.1016/j.apacoust.2010.05.009

Gillespie, D., and Caillat, M. (**2008**). "Statistical classification of odontocete clicks," Can. Acoust., **36**, 20–26. Retrieved from http://jcaa.caa-aca.ca/index.php/jcaa/article/view/1986

Gillespie, D., Dunn, C., Gordon, J., Claridge, D., Embling, C., and Boyd, I. (**2009**). "Field recordings of Gervais' beaked whales Mesoplodon europaeus from the Bahamas," J. Acoust. Soc. Am., **125**, 3428–33. doi:10.1121/1.3110832

Giorli, G., Goetz, K. T., Delarue, J., Maxner, E., Kowarski, K. A., Martin, S. B., Giorli, G., et al. (**2018**). "Unknown beaked whale echolocation signals recorded off eastern New Zealand," J. Acoust. Soc. Am., **285**, 1–8. doi:10.1121/1.5032127

Griffiths, E. T., Keating, J. L., Barlow, J., and Moore, J. E. (**2018**). "Description of a new beaked whale echolocation pulse type in the California Current," Mar. Mammal Sci., , doi: 10.1111/mms.12560. doi:10.1111/mms.12560

Hildebrand, J. A., Baumann-Pickering, S., Frasier, K. E., Trickey, J. S., Merkens, K. P., Wiggins, S. M., McDonald, M. A., et al. (**2015**). "Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico," Sci. Rep., **5**, 16343. doi:10.1038/srep16343

Houégnigan, L., Zaugg, S., van der Schaar, M., and André, M. (**2010**). "Space–time and hybrid algorithms for the passive acoustic localisation of sperm whales and vessels," Appl. Acoust., **71**, 1000–1010. doi:10.1016/j.apacoust.2010.05.017

Houser, D. S., Helweg, D. a, and Moore, P. W. (**1999**). "Classification of dolphin echolocation clicks by energy and frequency distributions," J. Acoust. Soc. Am., **106**, 1579–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10489713

Hsu, M. K., Sheu, J. C., and Hsue, C. (**2011**). "Overcoming the negative frequencies - Instantaneous frequency and amplitude estimation using Osculating Circle method," J. Mar. Sci. Technol., **19**, 514–521.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., et al. (**1998**). "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. R. Soc. A Math. Phys. Eng. Sci., **454**, 903–995. doi:10.1098/rspa.1998.0193

Ioup, G. E., Ioup, J. W., Pflug, L. A., Tashmukhambetov, A. M., and Sidorovskaia, N. A. (**2007**). "Acoustic identification of beaked and sperm whales," J. Acoust. Soc. Am., **122**, 3003. doi:10.1121/1.2942722

Jefferson, T. A., Leatherwood, S., and Webber, M. A. (**1993**). "FAO Species Identification Guide," Mar. Mamm. World, **13**, 326–339. doi:10.1643/0045-8511(2001)003[0212:]2.0.CO;2

Johnson, M., Madsen, P. T., Zimmer, W. M. X., de Soto, N. A., and Tyack, P. L. (**2004**). "Beaked whales echolocate on prey," Proc. Biol. Sci., **271**, S383–S386. doi:10.1098/rsbl.2004.0208

Johnson, M., Madsen, P. T., Zimmer, W. M. X., de Soto, N. A., and Tyack, P. L. (**2006**). "Foraging Blainville's beaked whales (Mesoplodon densirostris) produce distinct click types matched to different phases of echolocation," J. Exp. Biol., **209**, 5038–5050. doi:10.1242/jeb.02596

Kaiser, J. F. (**1990**). "On a simple algorithm to calculate the 'energy' of a signal," Int. Conf. Acoust. Speech, Signal Process., IEEE, 381–384. doi:10.1109/ICASSP.1990.115702

Kandia, V., and Stylianou, Y. (**2006**). "Detection of sperm whale clicks based on the Teager-Kaiser energy operator," Appl. Acoust., **67**, 1144–1163. doi:10.1016/j.apacoust.2006.05.007

Klay, J., Mellinger, D. K., Moretti, D. J., Martin, S. W., and Roch, M. A. (**2010**). "Advanced methods for passive acoustic detection, classification, and localization of marine mammals.,"

Klinck, H., and Mellinger, D. K. (**2011**). "The energy ratio mapping algorithm: a tool to improve the energy-based detection of odontocete echolocation clicks," J. Acoust. Soc. Am., **129**, 1807–12. doi:10.1121/1.3531924

Klinck, H., Mellinger, D. K., Klinck, K., Bogue, N. M., Luby, J. C., Jump, W. a., Shilling, G. B., et al. (**2012**). "Near-real-time acoustic monitoring of beaked whales and other cetaceans using a Seaglider," PLoS One, **7**, 1–8. doi:10.1371/journal.pone.0036128

LeBien, J. G., and Ioup, J. W. (**2018**). "Species-level classification of beaked whale echolocation signals detected in the northern Gulf of Mexico," J. Acoust. Soc. Am., **144**, 387–396. doi:10.1121/1.5047435

Lepper, P., Dümortier, N., Dudzinski, K., and Datta, S. (**2005**). "Separation of Complex Echolocation Signal 'Trains´ From Multiple Bio-Sonar Sources," Proc. Int. Conf. "Underwater Acoust. Meas. Technol. &Results" Heraklion, Crete, Greece, 28th June – 1st July 2005, 913–918.

MacLeod, C. D., Perrin, W., Pitman, R. L., Barlow, J., Ballance, L. T., D'Amico, A., Gerrodette, T., et al. (**2006**). "Known and inferred distributions of beaked whale species (Cetacea: Ziphiidae)," J. Cetacean Res. Manag., **7**, 271–286.

Madsen, P. T., Johnson, M. P., Aguilar de Soto, N., Zimmer, W. M. X., and Tyack, P. L. (**2005**). "Biosonar performance of foraging beaked whales (Mesoplodon densirostris)," J. Exp. Biol., **208**, 181–94. doi:10.1242/jeb.01327

Matsumoto, H., Jones, C., Klinck, H., Mellinger, D. K., Dziak, R. P., and Meinig, C. (**2013**). "Tracking beaked whales with a passive acoustic profiler float," J. Acoust. Soc. Am., **133**, 731–740. doi:10.1121/1.4773260

McDonald, M. A., Hildebrand, J. a, Wiggins, S. M., Johnston, D. W., and Polovina, J. J. (**2009**). "An acoustic survey of beaked whales at Cross Seamount near Hawaii," J. Acoust. Soc. Am., **125**, 624–7. doi:10.1121/1.3050317

Moore, J. E., and Barlow, J. P. (**2013**). "Declining abundance of beaked whales (family Ziphiidae) in the California Current large marine ecosystem," PLoS One, **8**, e52770. doi:10.1371/journal.pone.0052770

Nishiguchi, K., and Kobayashi, M. (**2000**). "Improved algorithm for estimating pulse repetition intervals," IEEE Trans. Aerosp. Electron. Syst., **36**, 407–421. doi:10.1109/7.845217

Podestà, M., D'Amico, a, Pavan, G., Drougas, a, Komnenou, a, and Portunato, N. (**2006**). "A review of Cuvier's beaked whale strandings in the Mediterranean Sea," J. Cetacean Res. Manag., **7**, 251–261.

Rankin, S., Baumann-Pickering, S., Yack, T., and Barlow, J. (**2011**). "Description of sounds recorded from Longman's beaked whale, Indopacetus pacificus," J. Acoust. Soc. Am., **130**, EL339-EL344. doi:10.1121/1.3646026

Roberts, B. L., and Read, A. J. (**2015**). "Field assessment of C-POD performance in detecting echolocation click trains of bottlenose dolphins (Tursiops truncatus)," Mar. Mammal Sci., **31**, 169–190. doi:10.1111/mms.12146

Roch, M. A., Batchelor, H., Baumann-Pickering, S., Berchok, C. L., Cholewiak, D., Fujioka, E., Garland, E. C., et al. (**2016**). "Management of acoustic metadata for bioacoustics," Ecol. Inform., **31**, 122–136. doi:10.1016/j.ecoinf.2015.12.002

Roch, M. A., Klinck, H., Baumann-Pickering, S., Mellinger, D. K., Qui, S., Soldevilla, M. S., and Hildebrand, J. A. (**2011**). "Classification of echolocation clicks from odontocetes in the Southern California Bight," J. Acoust. Soc. Am., **129**, 467–75. doi:10.1121/1.3514383

Roch, M. A., Soldevilla, M. S., Hoenigman, R., Wiggins, S. M., and Hildebrand, J. A. (**2008**). "Comparison of Machine Learning Techniques for the Classification of Echolocation Clicks from Three Species of Odontocetes," Can. Acoust., **36**, 41–47.

Roch, M. A., Stinner-sloan, J., Baumann-pickering, S., and Wiggins, S. M. (**2015**). "Compensating for the effects of site and equipment variation on delphinid species identification from their echolocation clicks," J. Acoust. Soc. Am., , doi: 10.1121/1.4904507. doi:10.1121/1.4904507

Rogers, T. L., and Brown, S. M. (**1999**). "Acoustic observations of Arnoux's beaked whale (Berardius arnuxii) off Kemp Land, Antarctica," Mar. Mammal Sci., **15**, 192–198.

Simard, P., Hibbard, A. L., McCallister, K. a, Frankel, A. S., Zeddies, D. G., Sisson, G. M., Gowans, S., et al. (**2010**). "Depth dependent variation of the echolocation pulse rate of bottlenose dolphins (Tursiops truncatus)," J. Acoust. Soc. Am., **127**, 568–78. doi:10.1121/1.3257202

Soldevilla, M. S., Henderson, E. E., Campbell, G. S., Wiggins, S. M., Hildebrand, J. A., and Roch, M. A. (**2008**). "Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks," J. Acoust. Soc. Am., **124**, 609–624. doi:10.1121/1.2932059

Starkhammar, J., Nilsson, J., Amundin, M., Kuczaj, S. A., Almqvist, M., and Persson, H. W. (**2011**). "Separating overlapping click trains originating from multiple individuals in echolocation recordings," J. Acoust. Soc. Am., **129**, 458–466. doi:10.1121/1.3519404

Torrey, L., and Shavlik, J. (**2009**). "Transfer learning," Handb. Res. Mach. Learn. Appl. Trends Algorithms, Methods Tech., Information Science Reference - Imprint of: IGI Publishing,

Hershey, PA, pp. 1–22. doi:10.1016/j.jbi.2011.04.009

Trickey, J. S., Baumann-Pickering, S., Hildebrand, J. a., Reyes Reyes, M. V., Melcón, M., and Iñíguez, M. (**2015**). "Antarctic beaked whale echolocation signals near South Scotia Ridge," Mar. Mammal Sci., **31**, 1265–1274. doi:10.1111/mms.12216

Tyack, P. L., Zimmer, W. M. X., Moretti, D., Southall, B. L., Claridge, D. E., Durban, J. W., Clark, C. W., et al. (**2011**). "Beaked whales respond to simulated and actual navy sonar," PLoS One, , doi: 10.1371/journal.pone.0017009. doi:10.1371/journal.pone.0017009

Venkitaraman, A., and Seelamantula, C. S. (**2012**). "A Technique to Compute Smooth Amplitude, Phase, and Frequency Modulations From the Analytic Signal," IEEE Signal Process. Lett., **19**, 623–626. doi:10.1109/LSP.2012.2209872

Wahlberg, M., Beedholm, K., Heerfordt, A., and Møhl, B. (**2011**). "Characteristics of biosonar signals from the northern bottlenose whale, Hyperoodon ampullatus," J. Acoust. Soc. Am., **130**, 3077–3084. doi:10.1121/1.3641434

Williams, R., Wright, A. J., Ashe, E., Blight, L. K., Bruintjes, R., Canessa, R., Clark, C. W., et al. (**2015**). "Impacts of anthropogenic noise on marine life: Publication patterns, new discoveries, and future directions in research and management," Ocean Coast. Manag., **115**, 17–24. doi:10.1016/j.ocecoaman.2015.05.021

Yack, T. M., Barlow, J., Calambokidis, J., Southall, B., and Coates, S. (**2013**). "Passive acoustic monitoring using a towed hydrophone array results in identification of a previously unknown beaked whale habitat," J. Acoust. Soc. Am., **134**, 2589–2595. doi:10.1121/1.4816585

Yack, T. M., Barlow, J., Roch, M. a., Klinck, H., Martin, S., Mellinger, D. K., and Gillespie, D. (**2010**). "Comparison of beaked whale detection algorithms," Appl. Acoust., **71**, 1043–1049. doi:10.1016/j.apacoust.2010.04.010

Zaugg, S., van der Schaar, M., Houégnigan, L., and André, M. (**2013**). "Extraction of pulse repetition intervals from sperm whale click trains for ocean acoustic data mining," J. Acoust. Soc. Am., **133**, 902–11. doi:10.1121/1.4773278

Zaugg, S., Schaar, M. Van Der, Houégnigan, L., and André, M. (**2012**). "A framework for the automated real-time detection of short tonal sounds from ocean observatories," Appl. Acoust., **73**, 281–290. doi:10.1016/j.apacoust.2011.09.009

Zaugg, S., van der Schaar, M., Houégnigan, L., Gervaise, C., and André, M. (**2010**). "Real-time acoustic classification of sperm whale clicks and shipping impulses from deep-sea observatories," Appl. Acoust., **71**, 1011–1019. doi:10.1016/j.apacoust.2010.05.005

Zimmer, W. M. X., Johnson, M. P., Madsen, P. T., and Tyack, P. L. (**2005**). "Echolocation clicks of free-ranging Cuvier's beaked whales (Ziphius cavirostris)," J. Acoust. Soc. Am., **117**, 3919. doi:10.1121/1.1910225

Chapter 2

DetEdit: A graphical user interface for annotating and editing events detected in acoustic data

2.1. Abstract

Many biological monitoring projects rely on the detection of species sounds to infer biologically and ecologically relevant information. Passive acoustic monitoring has become an increasingly important data collection method, providing massive datasets. However, these large datasets require advanced data processing techniques to make data analysis possible.

A MATLAB-based graphical user interface was developed for events detected in acoustic data, called *DetEdit*, to accelerate the process of editing and annotating extensive acoustic datasets. This tool is highly dynamic and can be used for multiple purposes, ranging from annotation of individual signals and evaluation of signal properties to removing false detections and obtaining false positive rates. *DetEdit* examines and displays different parameters of acoustic detections, including a time series of signal amplitude, long-term spectral averages, inter-pulse interval, and scatter plots of peak frequency, RMS amplitude, and peak-to-peak amplitude. Additionally, it displays either individual or averaged signal properties of waveforms and power spectra. With *DetEdit*, individual or sets of detections can be rapidly classified, annotated or removed.

This tool serves as a step to process acoustic data to create datasets of signals labeled to species-level or refine and evaluate detections for further analyses, such as automatic classification and density estimation. Although it has been designed for odontocete echolocation clicks, *DetEdit* can easily be adapted to almost any stereotyped signal. Our package complements currently available tools for the bioacoustic community and is provided open source with an extensive online user manual.

2.2. Introduction

A large variety of animals produce species-specific acoustic signals or calls, including marine mammals, fish, crustaceans, primates, bats, birds, anurans, and insects (Anorim, 2006; Brinkløv et al., 2013; Brown and Grinnell, 1980; Gerhardt et al., 2003; Hawkins, 1986; Kroodsma, 1982; Quam et al., 2017; Richardson et al., 1995; Versluis, 2000). Acoustic analysis has become a standard method in field studies of animal vocal communication, with studies generaly using manual-based approaches to detect acoustic cues in acoustic recordings. At a time when acoustic monitoring of populations generates a vast amount of data, the use of human analysis alone becomes unmanageable. Basic classification algorithms, or more complex methods such as machine learning, become a necessity for accelerating the analysis process. Automatic detectors provide more comparable estimates throughout the study period when processing long-term time series. They are less prone to bias from human analysts, and they can be quantified more objectively. However, the use of automatic detection algorithms is still cumbersome when preparing acoustic data for estimates of density of sound-producing animals from long-term acoustic recorders (Frasier, 2015; Harris et al., 2018; Hildebrand et al., 2015; Küsel et al., 2011; Marques et al., 2011). First, the design of algorithms that automatically identify target species is often beyond the capability of many ecologists. Second, labeling sounds to train automatic routines can be a prolonged process. Third, the detection algorithm performance must be estimated, which is mostly done manually by sampling a subset of the data or using some performance statistics (Marques et al., 2009). The measured performance is implied only for datasets which have been explicitly tested, and knowledge of the detector performance on other datasets for which density is being estimated is also required.

Analysis speed and misclassification of detections to species-level are two important factors in large-scale explorations of acoustic monitoring. The aim of this study is to increase the repertoire of tools available to make species assessment less time consuming and more efficient when processing long-term acoustic data. A new custom graphical user interface (GUI) based tool, *DetEdit*, is presented to accelerate and enhance the process of species-level analysis of acoustic data. This tool will facilitate the integration of advanced machine learning algorithms for species classification. It can be applied to stereotyped signals that are characterized in spectral shape, such as the underwater sounds produced by odontocetes, shrimp, lobsters, sonar, shipping, and echosounders, or terrestrial sounds, including the calls made by bats, oilbirds and swiftlets. Here the use of *DetEdit* is illustrated with two case studies of odontocetes from long-term time series.

2.3. Methods

DetEdit package provides a set of novel tools designed to parse, manipulate, and visualize acoustic detections in a workflow format using an interface tool developed in MATLAB (Mathworks, Natick, MA). The schema (**Figure 2.1**) illustrates ease of use and flexibility of this tool. The *DetEdit* processing pipeline begins by using acoustic detection times to build signal parameter matrices to then manually assess data with a sliding window approach, and modify parameter matrices based on the analyst decisions. This process is repeated as many times as needed. In addition, false detections can be manually assessed using the sliding window approach. If specified, acoustic detection parameters are illustrated in summary plots and stored.

The code repository includes all tools to organize and manage the data processing, and several examples to visualize and reproduce results. Two case studies are discussed in the following sections, describing the species classification algorithms implemented to obtain the acoustic cues and associated parameters used for editing detections with the *DetEdit* package.

2.3.1. Data analysis

2.3.1.1. Case Study 1: Sperm whales

A routine for identifying candidate sperm whale clicks implemented in MATLAB (Mathworks, Natick, MA) was developed to distinguish signals with a multi-step approach. Sperm whale wideband clicks can be easily distinguished from other marine species, but one of the difficulties is that their signals are highly similar to the impulsive cavitation sounds emitted by propellers of vessels. Potential click candidates were identified using a variation of the two-step approach proposed by Roch et al. (2011) and Soldevilla et al. (2008). The presence of an echolocation click was determined in the spectral frame (2048 FFT, 50% overlap, Hann window) when 20 % of the frequency bins between 5 – 20 kHz exceeded a signal-to-noise ratio -10dB threshold. The threshold was set low to allow detection of a large number of sperm whale clicks with the trade-off of many false positives, which were subsequently removed. In the second detection step, a high-resolution detector was used to return precise start and end times of individual impulses. All acoustic data were band passed (5 – 95 kHz) to minimize the influence of low-frequency noise from vessels, electrical noise, and weather. Impulses were identified within flagged times from the previous step using a Teager-Kaiser energy operator (Kaiser, 1990; Kandia and Stylianou, 2006). The noise floor was estimated at the 40^{th} percentile of the energy distribution, and Teager-Kaiser energy peaks that exceeded the noise floor by a factor of 7 were defined as potential echolocation clicks. Click start and end times were identified from each energy peak as the first and last points that were three times greater than the noise floor.

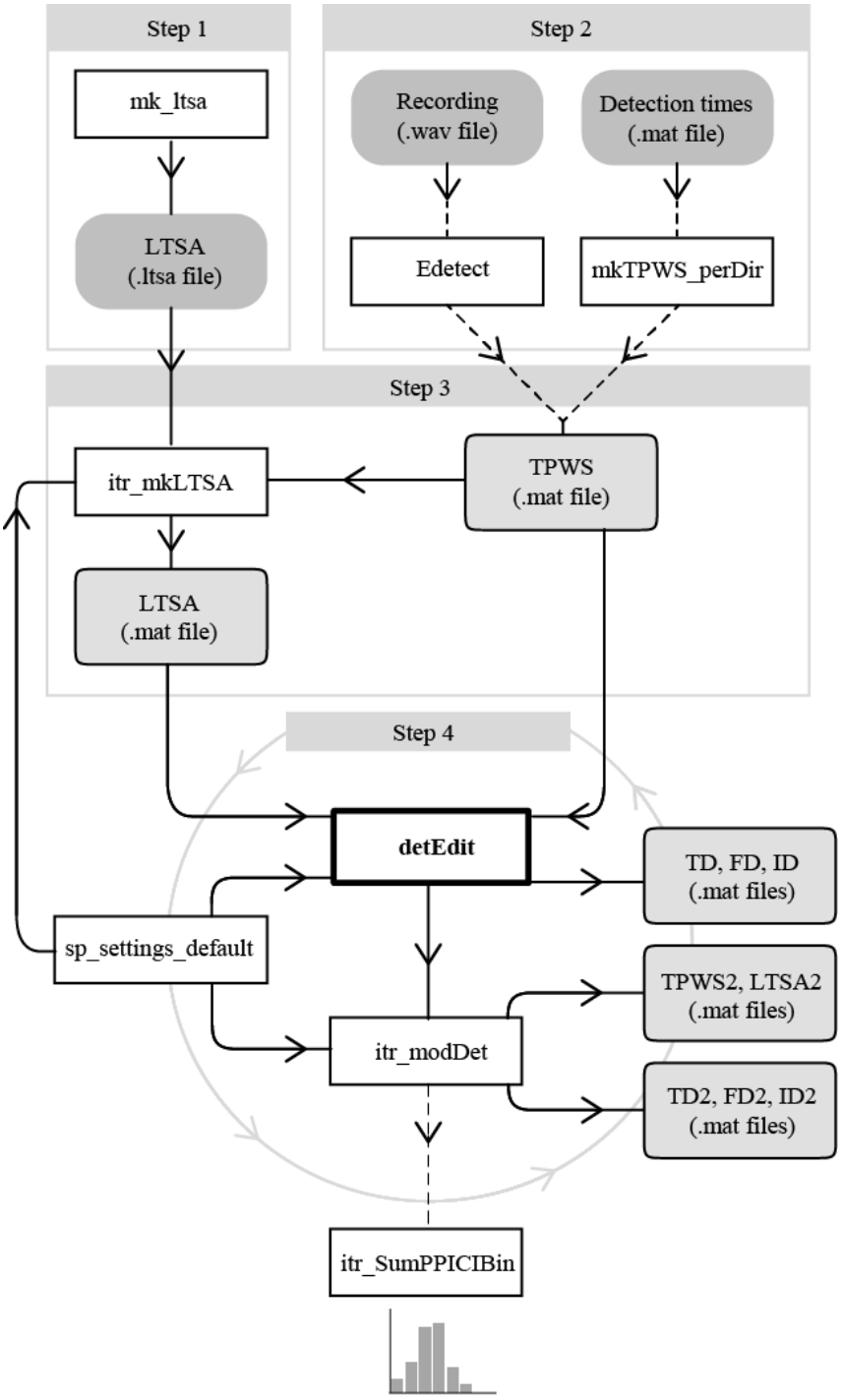


Figure 2.1. Schematic representation of the *DetEdit* workflow. To run the interface, the user (1) creates a Long-term spectral average (LTSA) of the data, (2) provides detections to create a TPWS (start Time, Peak-to-peak amplitude, Waveform, Spectra parameters) file or runs a simple energy detector, (3) creates LTSAs per bout, (4) runs the interface detEdit to manually edit detections, and (5) deletes false detections from data. If interested, histograms and plots are created from the final decisions. Software functions (white boxes), data files (gray boxes), dashed lines are optional steps, and solid lines indicate the data workflow.

Click times that were less than 30 ms apart were merged, considering that sperm whale clicks are characterized by multiple pulses that are nearly 5 ms apart (Zimmer et al., 2005). Spectra of each detected signal were calculated using 4 ms of data, and a 512-point Hann window centered on the click with 50 % overlap. Spectra were corrected for the frequency-dependent hydrophone response with a transfer function. To provide a consistent detection threshold, only clicks that exceeded a received level (RL) of at least 130 dB_{pp} re 1 µPa were retained for further analyses.

An automated classifier was developed to exclude periods of ship passages that were attributable to impulsive signals from the cavitation noise of ship propellers. The classification was used to identify ship passages from long-term spectral averages (LTSA, Wiggins and Hildebrand, (2007)) to remove detected impulses within the event. LTSAs are long-term spectrograms, which are created using a time average of 5 s and frequency bins of 100 Hz (calculated using Hanning window without overlap). Noise produced from the high-frequency acoustic recording packages (HARP) when writing to the disk was removed by excluding the first 15 s interval of each sequential 75 s acoustic record (Wiggins et al., 2016). Average power spectral density (APSD) estimates were then computed in 2-h blocks of data over three frequency bands, referred here and after as the low (1 – 5 kHz), medium (5 – 10 kHz), and high (10 – 50 kHz) band, and corrected for the hydrophone transfer function (**Figure 2.2**). Missing values due to the system noise removal were interpolated from neighboring averages. An adaptive threshold was used to determine transient signals (e.g., odontocetes, ship passages, weather) above the background noise over the three frequency bands. Two state levels of the ASPD were estimated using a histogram method, where the histogram was divided into two equal regions, and the mode of each region was taken as the lower and upper level. The time-dependent threshold was calculated as the mean of the lower- and higher-state levels. Start and end times of high amplitude events were identified when crossing over the threshold, and only those longer than 150 s were retained, referred here and after as transient.

Transients present in the 3 bands were classified as ship passages. However, cases of close encounters of odontocetes also showed transients in all three bands. These cases were identified as odontocetes only if the transient in the high band was longer than that in the medium band, and the duration of the transient in the medium band was 2/3 longer than that in the lower band. Presence of distant ships did not yield transients in the high band. Those passages were identified if duration in the low and medium band were above 250 s and differentiated from sperm whale encounters if transients in the medium band were at least 2/3 smaller than those in the lower band. Weather noise events (e.g., rain and storms) do not produce impulsive sounds. These events were not excluded from analysis and were distinguished from the ship passages by comparing RLs. RLs were taken from 1-min sliding-windows with 55 s overlap of the APSD estimates. RLs of the transients in the three bands were compared, and ship passages were defined if:

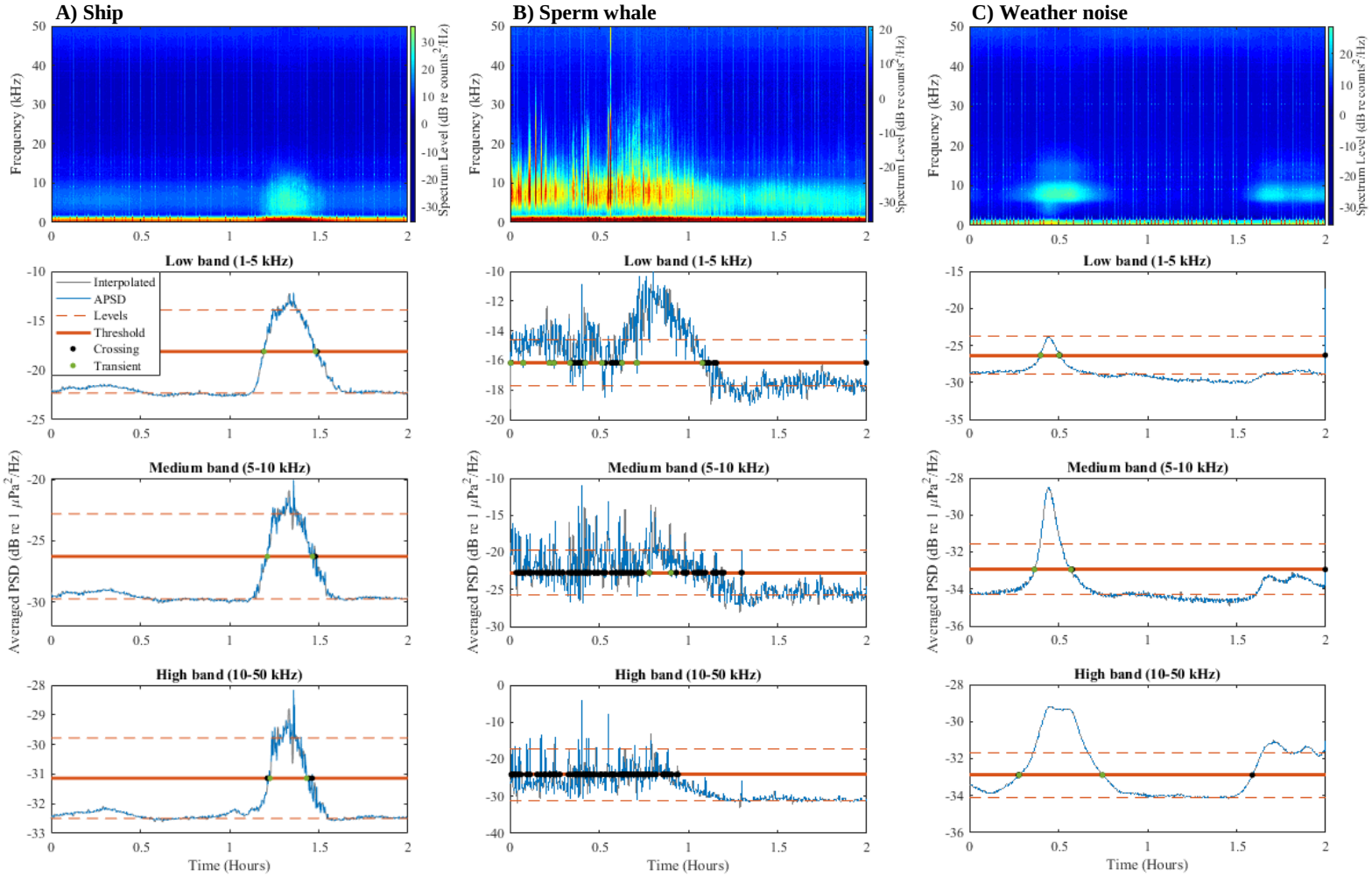


Figure 2.2. Event types including (A) a ship passage, (B) sperm whale encounter, and (C) weather noise in 2-h LTSAs (upper panels). LTSAs are 2000 samples (100 Hz bins) with 50% overlap and color represents sound pressure spectrum level. Concurrent averaged power spectral densities (APSD) for the three frequency bands (lower panels) with transients identified if duration between start and end crossing points of the event was above 150 s.

- RLs in the low band were above 80 dB, while below 60 dB in the medium and higher band or,
- The difference between RLs in the low band and the mean of RLs in the medium and higher band were above 15.

To ensure that ship passages on the edges of the window were also detected, an overlapping sliding window classification method was implemented. Using 2-h window with 30 min overlap before and after, the decision of ship presence was determined if ship passages were detected in two of the three overlapping windows. All identified ship passages were then manually revised by an experienced acoustician using the custom software program *Triton* (Wiggins and Hildebrand, 2007) to minimize the inclusion of such events in the sperm whale dataset. False positives were a result of misclassification between weather events, and ship passages, only a misclassification rate below 0.004 was obtained for sperm whale encounters.

A basic classification method based on spectral click shape was implemented to discard obvious non-sperm whale clicks, such as beaked whale and delphinid clicks. Often peak frequencies were used to discriminate sperm whale clicks because of their lower frequency compared to other odontocetes. When animals were very close to the acoustic recorder clicks often had more energy in the higher frequencies; thus, making it more difficult to differentiate them with confidence. Clicks were only removed from the analysis if they had a peak frequency above 50 kHz or a second peak above 22 kHz. Sperm whales have a distinctive spectral shape with most energy in frequencies below 30 kHz, while delphinids and beaked whales have more energy in higher frequencies; however, spectral shapes for dolphins and beaked whales are expected to have more variation because of the effect of acoustic attenuation for high frequencies. To characterize the spectral shape of each click, the difference across spectral levels of a specified frequency band was computed as

$$\Delta u = mean(u_{high}) - mean(u_{low})$$

where u_{high} is the vector of spectral levels of one click across the frequency range of 35 – 45 kHz, and u_{low} of 10 – 20 kHz. Click spectra were normalized between [0, 1] as

$$u_n = \frac{u - \min(u)}{\max(u - \min(u))}$$

where u is the vector of spectral levels of one click across the frequency range of interest, and u_n is the normalized spectral level of that click. The median of the normalized spectra above 50 kHz was computed. Signals with a spectral difference (Δu) above -5 and a median of the normalized spectra (u_n) above 0.8 were defined as high-frequency spectral shapes, which were linked to beaked whale and delphinid clicks. If more than 85 % of the clicks were discarded in the audio file, the section was linked to a specific anticipated false positive source (e.g. ship noise, seismic

surveys, delphinid, and beaked whale) and was discarded from further analysis. Isolated clicks separated from neighboring signals by more than 5 seconds were also discarded.

2.3.1.2. Case Study 2: Dolphins

Frasier et al. (2017) developed a workflow to automatically identify distinct dolphin click types within large datasets using unsupervised clustering tools. This method was used to automatically classify signals into categories and to assist human analysts with processing multi-species acoustic encounters. Only clicks that exceeded a RL of at least 120 dB_{pp} re 1 µPa were included.

The content of both acoustic datasets and details of the methods involving *DetEdit* are discussed and illustrated with examples in **Section 2.4.1.2**.

2.3.2. Design and Implementation

The *DetEdit* infrastructure and typical workflow are summarized in **Figure 2.1**. The package, which depends on the core `detEdit`, implements a hierarchical automated pipeline that incorporates data preprocessing, visualization, and manipulation tools. **Table 2.1** lists the main functions found in *DetEdit* with a brief summary of the usage. Some tools (`Edetect`, `Edetect_wav`, `mkTPWS_perDir`) are pre-processing pipelines, taking a vector of detections or applying a generic detector to produce matrices of specific parameters of acoustic detections. Others (`mk_ltsa`, `mkLTSAsession`), compute spectral averages of data. The primary tool, or the interface tool `detEdit`, is able to complete a range of analyses starting with visualization of the data with specific parameters, interactively exploring plotted data, and manually annotating data. Others, such as `modDet`, and `SumPPICIBin` are post-processing pipelines, taking user annotated decisions, manipulating data detections, and producing different types of plot summaries.

2.3.2.1. Features

The graphical user interface allows users to intuitively and efficiently annotate data using seven different panels revealing high discrimination abilities. These panels display peak-to-peak amplitudes, LTSA, inter-detection-intervals (IDI), waveform, spectra, root-mean-square (RMS) and peak frequencies in an interactive plotting mode to allow users to manually edit large samples of detections.

Peak-to-peak amplitudes of detected clicks are displayed per bout (**Figure 2.3**), with the concurrent LTSA and IDIs. The IDI is computed as the difference between one impulse and the next. This parameter is variable because if there are multiple animals recorded, the received impulses will be interleaved with each other. As a result, even if there is a single animal there could be variability in IDI due to detectability when the animal changes direction. However, there is an undeniable consistency in IDI. Fundamentally, when animals start to approach the sensor while

clicking, because their signals are directional, on-axis signals are detected. As soon as they go past or point away from the sensor, the off-axis signals are detected. When a group of animals are at some distance from the sensor and are scanning with the signals, only the on-axis signals are detected. Hence, at any given time when an animal happens to be pointing at the sensor and clicking, its characteristic cue rate (a consistent IDI) is received. The IDI, together with peak-to-peak amplitudes of the received signals through the encounter, and the concurrent spectral characteristics distinguishable with the LTSA panel allow interpretation of the data and correct differentiation of signals. Usually animals are detected from a distance, but when closer to the sensor, multiple animals and off-axis signals are received thus showing higher peak amplitudes and IDIs that are less than the characteristic rate for that species.

From the prospective of peak-to-peak received level, two signals can have the same values, but have different spectral content and duration. Displaying peak-to-peak amplitudes with respect to RMS received levels allows pulses that would have been received with the same peak-to-peak level to be distinguished from those that are unrelated. In the interface, these parameters are represented with a transformation in the slope, shifted by a specified value. This allows shifting the slope correlation to be vertical for easing the annotation of detections. In a separate window peak frequency with respect to RMS are displayed following the same slope shift from the peak-to-peak amplitudes with RMS, adding discrimination capabilities in the spectral content of the signals. Another window shows averaged spectra and waveforms of all signals within bouts, and individual spectra or waveforms of selected detections can be compared. The characteristic peaks in the spectra and the duration of a signal are powerful parameters to distinguish species, hence these two parameters are included in the interface.

2.3.2.2. Workflow

The process involves following several steps to create the parameters needed for the panels of the interface (**Figure 2.1**, **Table 2.1**):

Step 1: Create LTSA files. The package relies on `mk_ltsa` function to read audio files (wave format) and compress data into long-term spectral averages (LTSA) that allows exploration of long-term acoustic data. An LTSA of a time-series data is essentially a long-term spectrogram and is produced for selected audio files by specifying the time average length and frequency bin size to compress the data.

Step 2: Create TPWS files. The input to the *DetEdit* GUI is a MATLAB file labeled with "TPWS" (start Time, Peak-to-peak amplitude, Waveform and Spectra parameters) that contains matrices of the parameters from acoustic detections. These matrices can be created manually by the user or with the `mkTPWS_perDir` function that builds the four primary variables to visualize detections in the interface. These variables include:

- `MTT`: a vector of start times of detections.
- `MPP`: a vector of peak-to-peak received level amplitudes.
- `MSP`: a matrix (NxF) of detection spectra, where F is dictated by the size of the FFT used to generate the spectra and N, the number of detections.
- `MSN`: a matrix (NxF) of waveforms, where F is dictated by the window length used to extract the time series and N, the number of detections.

If no detections are given, the package provides the `Edetect` function to assist users in detecting acoustic events. This generic detector is based on a received level threshold that detects high-amplitude signals of a time series above the specified threshold.

Step 3: Create LTSA files per encounter. Detections are grouped in time bouts defined by the user. Click bouts are defined as periods of clicking with more than a specific time gap without detections both before and after. Bouts whose duration is shorter than a specified minimum length, and bouts containing fewer than a specified number of clicks are excluded from analysis. The `itr_mkLTSA` function takes an LTSA and produces two variables needed to represent LTSA per bouts in the interface. These variables are:

- `pt`: a vector of start times of spectral averages.
- `pwr`: a matrix (NxF) of power spectral densities, where F is dictated by the step size used to average frequency spectrum bins.

Visualization, annotation, and manipulation of acoustic data. After building the parameter files, the user evokes the interface by calling the `detEdit` function after specifying the directories and specific parameters in `detEdit_settings` function. Predefined parameters for eleven species of odontocetes (e.g. beaked whale, dolphin, sperm whale, narwhal and beluga) are provided in `sp_setting_default`. Seven windows are displayed to assess the acoustic detections (**Figure 2.3**, **Figure 2.3**. Example of annotating sperm whale detections with *DetEdit*. The `detEdit` function displays the graphical user interface (GUI) that allows users to annotate detections from continuous recordings organized in bouts of detections. Seven plots are displayed, (A) peak-to-peak amplitudes over time, (B) LTSA over time, (C) inter-click-intervals over time, (D) averaged spectra, (E) averaged waveform, (F) Peak-to-peak amplitudes over RMS, and (G) Peak frequencies over RMS. True detections shown in blue and manually identified false detections in red, delphinid clicks in this case. Manual inspection of one detection with the selection tool is displayed in black. Thresholds are displayed as a continuous red line (F and G).). The interactive visualization supports exploration of detections and their discriminative features. Following a simple list of keyword shortcut commands and the use of a paintbrush tool (specified on the user's manual at https://github.com/ScrippsWhaleAcoustics/DetEdit/wiki/How-to-use-it), the user can parse the data by manually selecting single or multiple detections to interactively visualize the

features or averaged features of the selection, and compare the parameters with the remaining detections within the displayed bout. **Figure 2.3** shows an example of editing acoustic data with sperm whale echolocation clicks within a mixed species time period. Dolphin clicks are distinguished from sperm whale clicks by comparing peak-to-peak amplitude and peak frequencies with the RMS. In this case, all delphinid detections are labeled as false detections (shown in red). Individual selection of clicks is shown in black. All detections are shown in gray on the background to ease the comparison. Thresholds can be defined for RMS, peak frequency, and peak-to-peak amplitudes to automate the process of annotating data. Therefore, any detection below the selected thresholds will be displayed automatically as false positive detections. But all annotations can be reversed and labeled as true positives or specified detection types by using a palette of colors with the Matlab paintbrush tool. All the changes made during the interactive interface are updated and stored in the corresponding files. False detection files (FD.mat file format) contain all start times from the TPWS files defined as false detections. Detection type files (ID.mat file format) contain all start times defined as a specified type of detection, with the corresponding color which defines the species or label type. **Figure 2.4** shows an example of using predefined detection settings in delphinid encounters, which ease the visualization of multispecies encounters and the processing of multi-species encounters. Detections can also be labeled as misidentifications and are stored in the misidentified files (MD.mat file format).

Post-processing. After manually editing all detections, decisions stored in different file types are used to modify the detection parameters stored in the TPWS files. The `itr_modDet` function excludes all the false detections and ID detections (if specified from the user) from the TPWS files. Apart from excluding the false detections, users can obtain exploratory plots of different features for each file. The process of using `detEdit` and `modDet` can be repeated iteratively until all detections are labeled, or a sufficiently low percentage of false detections is obtained (**Figure 2.1**). For every iteration, all files are stored together in a common directory to keep track of all changes.

Our goal was to develop the essential tools to process the data as efficiently as possible and to characterize the detectors for a density estimation framework. A post-processing tool was added into *DetEdit* to allow the characterization of the detector's false positive rate by inspecting random bouts with the interface. When specified, the interface displays random detections, and the user must decide if the detection is correct or false. A matrix with the false positive rate is built and stored in true detection files (TD.mat file format). **Figure 2.5** shows an example of using the false positive estimation procedure, which can be done at a signal level decision or by groups of signals within a time-bin interval. Only the procedure per signals is shown here, where one individual signal is displayed at a time within the encounter and defined as false or true following a simple keyword shortcut command. The time-bin interval procedure consists of deciding if at least one

signals within the defined interval is true. Signals or intervals evaluated for false positive rate estimation are selected randomly without replacement. Finally, a summary of all the data per site is given with histogram plots of peak-to-peak amplitudes, IDIs, and peak frequencies. In addition, a time series of daily and weekly presence of the true detections is represented.

Table 2.1. DetEdit function list and summary.

GENERAL FUNCTIONS	
Name	**Description**
`detEdit`	Displays interface and returns user decisions in files grouping the click parameters in *FD.mat, *TD.mat, *ID.mat, and *MD.mat files.
`detEdit_settings`	Setup specific parameters to run `detEdit`. Parameters can be specified to overwrite species default settings defined in `sp_settings_defaults`.
`Edetect`	Simple energy threshold detector to detect signals from time series. If desired can be applied per audio file using `Edetect_wav`.
`itr_mkLTSA`	Used to iterate between specified files to run `mkLTSAsessions`.
`itr_modDet`	Used to iterate between specified files to run `modeDet`.
`itr_SumPPICIBin`	Used to iterate between specified files to run `SumPPICIBin`.
`mkLTSAsessions`	Makes LTSAs per bout from LTSA of the corresponding deployment. Returns vector containing start times of spectral averages and power spectral densities into *LTSA.mat files.
`mkTPWS_perDir`	Groups detections in bouts and returns click parameters into *TPWS.mat files.
`modDet`	Deletes false detections stored in *FD.mat files. If given the option, it deletes ID detections stored in *ID.mat files.
`sp_settings_defaults`	Setup primary parameter for specific species. Eleven species of odontocetes are predefined.
`SumPPICIBin`	Plots peak-to-peak amplitude, peak frequency, inter-click-interval histograms, and daily and weekly presence. Computes parameters for density estimation.
DATA FILES	
FD.mat	Vector of start times of false detections.
ID.mat	Matrix of start times of detection types and ID number of detection type.
LTSA.mat	LTSA parameters per bout. Vector of start times of spectral averages (pt), and matrix of power spectral densities (pwr).
MD.mat	Vector of start times of misidentified detections.
TD.mat	Matrix of start times of true detections and user decision of the false positive rate analysis.
TPWS.mat	Parameters of detected signals. Vector of start times (MTT), vector of peak-to-peak amplitudes (MPP), matrix of times series (MSN), and matrix of spectrums (MSP).

Additional information about the optional parameters and their default settings can be found by viewing the HTML function help pages located at https://github.com/ScrippsWhaleAcoustics/DetEdit/wiki/How-to-use-it.

2.4. Results

Since the initial development of *DetEdit*, the package has been used to analyze a number of acoustic recordings (e.g., DCLDE2018 workshop dataset; Frasier et al., 2017; Hildebrand et al., 2015), and continues to be used for ongoing work. Here, the utility of *DetEdit* for facilitating analyses of large datasets is demonstrated with two case studies that used the integrated pipeline to perform species-level analyses of acoustic data. These examples are included in the package (from default species setting parameters to actual parameter files) reproducing results for a variety of encounters and conditions. Data were grouped in bouts of detections 30 minutes apart. Bouts of less than 75 s duration were discarded in both case study examples.

2.4.1.1. Case Study 1: Sperm whales

This first case study shows how *DetEdit* can be used to manually label false detections and its versatility for different labeling purposes. Data from three different sites in the Gulf of Mexico between 2010 and 2017 were processed using this interface. A total of 6,438 h of data (202 TB) were analyzed, and 34 million sperm whale clicks were verified and corresponding false positive rates were calculated.

The window panels of both peak-to-peak amplitude and peak frequencies over RMS were particularly useful in distinguishing sperm whales from other odontocetes, i.e. delphinid and beaked whale clicks. As shown in **Figure 2.3**, the removal of clicks from a dolphin encounter was possible by selecting those clicks that had a lower RMS. To accelerate the removal of low RMS clicks, an RMS threshold of 95 dB was implemented to automatically label all detections below the threshold as false positives (**Figure 2.3F**). The straight-line threshold was possible to implement because data plotting of RMS was transformed by a slope of 1.05, specified on the species settings.

Remaining impulses of ship passages that were not excluded with the ship detector, were mostly identified using the IDI plot and the spectral characteristics in the LTSA. When ships where present, IDI was not consistent and were labeled with the color code tool to identify those periods. These times were stored in the detection type files (ID.mat files) allowing exclusion as “no effort” times instead of false positives (which are stored in the FD.mat files) for congruence with the classifier algorithm process.

2.4.1.2. Case Study 2: Dolphins

Data from three different sites in the Gulf of Mexico between 2010 and 2016 were processed to detect echolocation clicks of different species of dolphins using unsupervised clustering. A total of 5,446 h of data (171 TB) were analyzed, and 115 million dolphin clicks were verified, and corresponding false positive rates for each species were calculated. Clusters were evaluated using this interface, displayed in different color codes to distinguish the multiple detection types

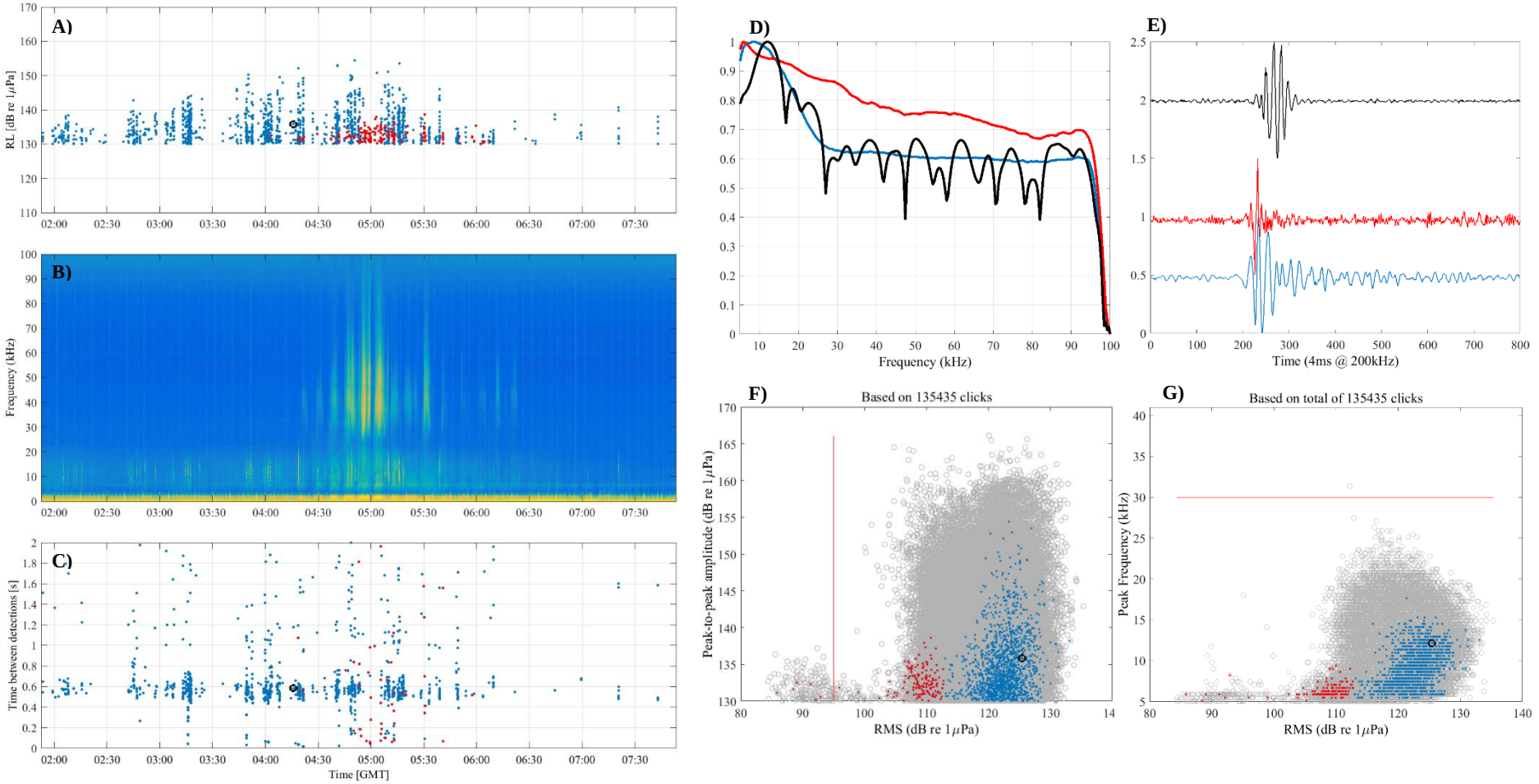


Figure 2.3. Example of annotating sperm whale detections with *DetEdit*. The `detEdit` function displays the graphical user interface (GUI) that allows users to annotate detections from continuous recordings organized in bouts of detections. Seven plots are displayed, (A) peak-to-peak amplitudes over time, (B) LTSA over time, (C) inter-click-intervals over time, (D) averaged spectra, (E) averaged waveform, (F) Peak-to-peak amplitudes over RMS, and (G) Peak frequencies over RMS. True detections shown in blue and manually identified false detections in red, delphinid clicks in this case. Manual inspection of one detection with the selection tool is displayed in black. Thresholds are displayed as a continuous red line (F and G).

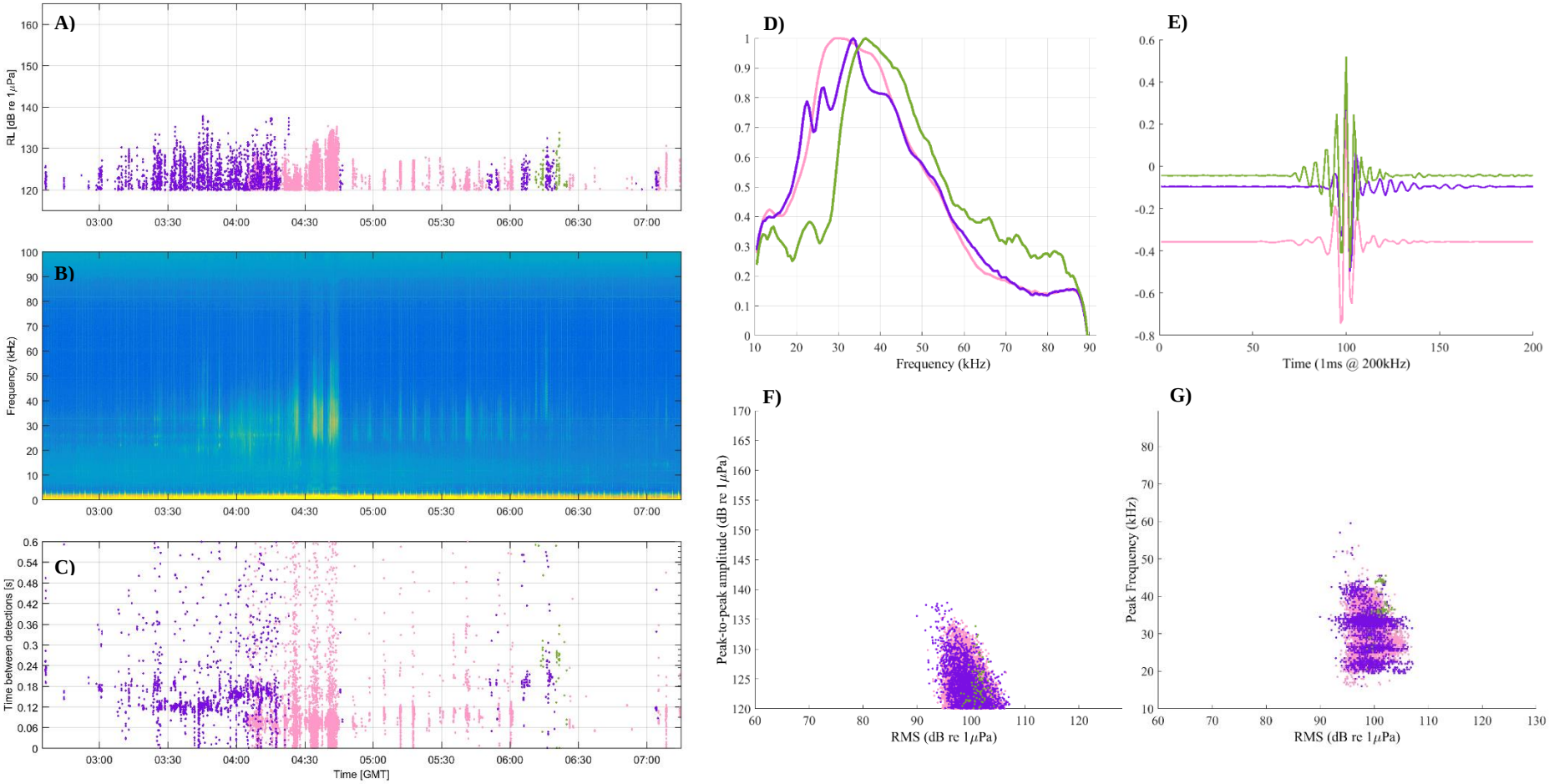


Figure 2.4. Example of annotated delphinid and beaked whale detections with *DetEdit.* The `detEdit` function displays the graphical user interface (GUI) that allows users to annotate detections from continuous events organized in bouts of detections. Seven plots are displayed, (A) peak-to-peak amplitudes over time, (B) LTSA over time, (C) inter-click-intervals over time, (D) averaged spectra, (E) averaged waveform, (F) Peak-to-peak amplitudes over RMS, and (G) Peak frequencies over RMS. ID click types represented in different colors, *Stenella spp*. in pink, Risso's dolphin in purple, and Gervais' beaked whale in green.

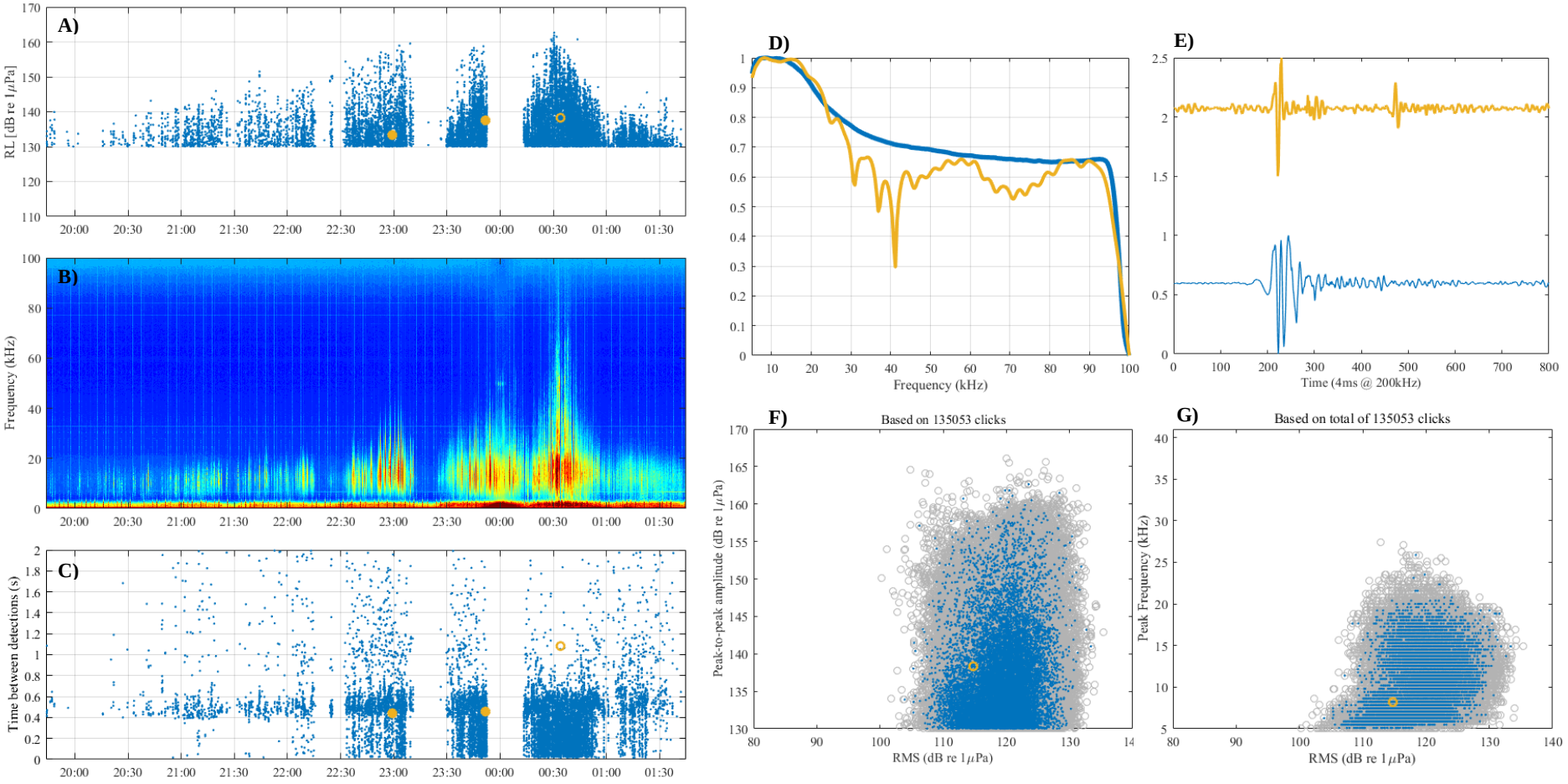


Figure 2.5. Example of evaluating false positives of sperm whale detections with *DetEdit*. The `detEdit` function displays the graphical user interface (GUI) that allows users to annotate detections from continuous events organized in bouts of detections. Seven plots are displayed, (A) peak-to-peak amplitudes over time, (B) LTSA over time, (C) inter-click-intervals over time, (D) averaged spectra, (E) averaged waveform, (F) Peak-to-peak amplitudes over RMS, and (G) Peak frequencies over RMS. Detected signals being evaluated within the encounter are shown in yellow. Evaluation is done in a consecutive manner, with the current signal marked with a yellow circle, and previously evaluated signals displayed in yellow dots here to ease the visualization. here to ease the visualization.

identified. As shown in **Figure 2.4**, multiple encounters of overlapping species were visually distinguishable and the ability to flag clicks in different colors made identification of different species possible. The selection tool supported this process by allowing manual selection of individual or multiple detections to compare with the different color-coded detections. The removal of non-delphinid clicks was possible by inspecting the peak-to-peak amplitude and peak frequency over RMS plots. IDI plots showed different consistent peaks in the distribution of IDIs, making it easy to identify the corresponding species **Figure 2.4C**. Furthermore, patterns in the received level window can also be indicative of different odontocetes signals. Dolphins typically have less directional, lower-amplitude beams than beaked whales and sperm whales, due to their smaller size. Received levels in dolphin echolocation clicks encounters tend to ramp up more slowly and reach lower peak amplitudes than beaked whale encounters which tend to be brief and higher amplitude, limited to the period when an animal is on or nearly on-axis.

2.5. Availability and future directions

DetEdit is available as a MATLAB package, with example datasets for different species on GitHub at the following link: https://github.com/ScrippsWhaleAcoustics/DetEdit. Detailed instructions are hosted on the Wiki section on GitHub and in a README file included in the package. The package was first developed using MATLAB R2013b; future work of *DetEdit* will focus on integrating the package on newer versions of MATLAB.

The genesis of this project grew from the day-to-day work and the need to address gaps between the tools provided and the analysis needs. The first significant contribution provided by *DetEdit* is the ability to accelerate the analysis process of acoustic data and removal of false positive within bouts of true detections. The second highly useful contribution is the ease of calculating false positive rates. *DetEdit* provides a way for analysts to easily use these tools and build custom workflows within the package.

Additional improvements could be made to increase the capabilities of species-level analyses of acoustic data. For example, additional clustering techniques could be used to visualize template clusters of target signals and provide certainty scores to assist labeling or removing detections.

2.6. Acknowledgments

Support for the development of DetEdit was provided by The Center for Integrated Modeling and Analysis of the Gulf Ecosystem (CIMAGE) Consortium of the BP/Gulf of Mexico Research Initiative (SA 12-10/GoMRI-007). I thank Steve Murawski and Sheryl Gilbert of University of South Florida (USF) for project assistance.

2.7. References

Anorim, M. C. P. (**2006**). "Diversity of sound production in fish," In B. G. Collin, S.P., Moller, P., Kapoor (Ed.), Commun. Fishes, Science Publishers, Enfield, HH, pp. 71–105.

Brinkløv, S., Fenton, M. B., and Ratcliffe, J. M. (**2013**). "Echolocation in oilbirds and swiftlets," Front. Physiol., **4**, 1–12. doi:10.3389/fphys.2013.00123

Brown, P. E., and Grinnell, A. D. (**1980**). "Echolocation Ontogeny in Bats," In R.-G. Busnel and J. F. Fish (Eds.), Anim. Sonar Syst., Springer US, Boston, MA, pp. 355–377. doi:10.1007/978-1-4684-7254-7_15

Frasier, K. E. (**2015**). *Density estimation of delphinids using passive acoustics: A case study in the Gulf of Mexico*. Ph.D. Thesis, University of California San Diego, La Jolla, CA, 262 pages.

Frasier, K. E., Roch, M. A., Soldevilla, M. S., Wiggins, S. M., Garrison, L. P., and Hildebrand, J. A. (**2017**). "Automated classification of dolphin echolocation click types from the Gulf of Mexico," PLOS Comput. Biol., **13**, e1005823. doi:10.1371/journal.pcbi.1005823

Gerhardt, H. C., Huber, F., and Simmons, A. M. (**2003**). "*Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions*," J. Acoust. Soc. Am., , doi: 10.1121/1.1591773. doi:10.1121/1.1591773

Harris, D. V., Miksis-Olds, J. L., Vernon, J. A., and Thomas, L. (**2018**). "Fin whale density and distribution estimation using acoustic bearings derived from sparse arrays," J. Acoust. Soc. Am., **143**, 2980–2993. doi:10.1121/1.5031111

Hawkins, A. D. (**1986**). "Underwater Sound and Fish Behaviour," In T. J. Pitcher (Ed.), Behav. Teleost Fishes, Springer US, Boston, MA, pp. 114–151. doi:10.1007/978-1-4684-8261-4_5

Hildebrand, J. A., Baumann-Pickering, S., Frasier, K. E., Trickey, J. S., Merkens, K. P., Wiggins, S. M., McDonald, M. A., et al. (**2015**). "Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico," Sci. Rep., **5**, 16343. doi:10.1038/srep16343

Kaiser, J. F. (**1990**). "On a simple algorithm to calculate the 'energy' of a signal," Int. Conf. Acoust. Speech, Signal Process., IEEE, 381–384. doi:10.1109/ICASSP.1990.115702

Kandia, V., and Stylianou, Y. (**2006**). "Detection of sperm whale clicks based on the Teager-Kaiser energy operator," Appl. Acoust., **67**, 1144–1163. doi:10.1016/j.apacoust.2006.05.007

Kroodsma, E. H. (**1982**). *Acoustic Communication in Birds: Production, Perception and Design Features of Sounds*, (H. O. Kroodsma, Edward H. Miller, Ed.) Acoustic Communication in Birds, Elsevier Science, 360 pages.

Küsel, E. T., Mellinger, D. K., Thomas, L., Marques, T. a, Moretti, D., and Ward, J. (**2011**). "Cetacean population density estimation from single fixed sensors using passive acoustics," J. Acoust. Soc. Am., **129**, 3610–3622. doi:10.1121/1.3583504

Marques, T. a., Munger, L., Thomas, L., Wiggins, S., and Hildebrand, J. a. (**2011**). "Estimating north pacific right whale Eubalaena japonica density using passive acoustic cue counting," Endanger. Species Res., **13**, 163–172. doi:10.3354/esr00325

Marques, T. a, Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. L. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales," J. Acoust. Soc. Am., **125**, 1982–94. doi:10.1121/1.3089590

Quam, R. M., Ramsier, M. A., and Fay, R. R. (**2017**). *Primate Hearing and Communication*, (A. N. Popper, Ed.) Springer Handbook of Auditory Research, Springer International Publishing, Cham, Vol. 63. doi:10.1007/978-3-319-59478-1

Richardson, W. J., Greene, C. R. J., Malme, C. I., and Thomson, D. H. (**1995**). *Marine Mammals and Noise*, Academic Press, San Diego, 576 pages.

Roch, M. A., Klinck, H., Baumann-Pickering, S., Mellinger, D. K., Qui, S., Soldevilla, M. S., and Hildebrand, J. A. (**2011**). "Classification of echolocation clicks from odontocetes in the Southern California Bight," J. Acoust. Soc. Am., **129**, 467–75. doi:10.1121/1.3514383

Soldevilla, M. S., Henderson, E. E., Campbell, G. S., Wiggins, S. M., Hildebrand, J. A., and Roch, M. A. (**2008**). "Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks," J. Acoust. Soc. Am., **124**, 609–624. doi:10.1121/1.2932059

Versluis, M. (**2000**). "How Snapping Shrimp Snap: Through Cavitating Bubbles," Science (80-.)., **289**, 2114–2117. doi:10.1126/science.289.5487.2114

Wiggins, S. M., Hall, J. M., Thayre, B. J., and Hildebrand, J. A. (**2016**). "Gulf of Mexico low-frequency ocean soundscape impacted by airguns," J. Acoust. Soc. Am., **140**, 176–183. doi:10.1121/1.4955300

Wiggins, S. M., and Hildebrand, J. A. (**2007**). "High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring," Int. Symp. Underw. Technol. 2007 Int. Work. Sci. Use Submar. Cables Relat. Technol. 2007, Institute of Electrical and Electrongics Engineers, Tokyo, Japan, 551–557.

Zimmer, W. M. X., Madsen, P. T., Teloni, V., and Tyack., M. P. J. and P. L. (**2005**). "Off-axis effects on the multipulse structure of sperm whale usual clicks with implications for sound production," J. Acoust. Soc. Am. 118(5)3337-3345. 2005., **118**, 9. doi:10.1121/1.2082707

Chapter 3

Population structure and patterns of habitat use of sperm whales in the Gulf of Mexico revealed by seven years of passive acoustic monitoring

3.1. Abstract

The population structure and seasonal movements of sperm whales (*Physeter macrocephalus*) in the Gulf of Mexico is poorly understood. Presence of sperm whale regular clicks were detected using passive acoustic monitoring along the continental slope in the Gulf of Mexico at three sites during and following the *Deepwater Horizon* oil spill (2010-2017). A method was developed to categorize acoustic encounters of sperm whales using the distribution of distinct inter-click intervals as potential indicators of sperm whale sex and population structure. These classes were determined by sub-sampling click series and correlating the mean inter-click interval (ICI) and the mean inter-pulse-interval (IPI) within each series. The inter-pulse intervals were then converted into size categories using Gordon's method for estimating animal acoustic length and further correlated to the inter-click intervals. Three different classes were found, one with a mean ICI and IPI of 0.8 s and 4.6 ms respectively, which corresponds to adult males, a second with mean ICI and IPI of 0.6 s and 3.6 ms respectively which corresponds to social units (mixed groups) of sperm whales (adult females and their offspring), and a third class with mean ICI and IPI of 0.7 s and 4 ms respectively believed to contain adult females and sub-adult males. The daily presence and seasonal variability of the three classes were evaluated. The mixed group was present all year at the northern sites of the Gulf but were only seasonally present at the southern site. Adult males were occasionally present throughout the year at two sites including, the female's core area near the Mississippi Canyon and at the eastern site. The seasonal presence of adult males may be related to seasonal breeding, in accordance with what has been previously observed in other populations in the South Pacific and in the Gulf of California.

3.2. Introduction

Sperm whales are listed globally as an endangered species under the Endangered Species Act (ESA) mainly due to commercial whaling which ended in 1988 with the implementation of the whaling moratorium by the International Whaling Commission (IWC). Today sperm whales face additional threats, in particular, anthropogenic noise and cumulative risks from multiple stressors including ship strikes, fisheries interactions, oil spills, and pollution (National Academies of Sciences Engineering and Medicine, 2017). The *Deepwater Horizon* (DWH) oil spill (NRDA, 2016) in the Gulf of Mexico (GOM) provided an impetus for developing a better understanding of sperm whale behavior and spatial use of the Gulf waters (Jochens et al., 2008).

Sperm whales are the most sexually dimorphic cetacean and despite being widely distributed (Rice, 1989), the sexes also have stratified distributions and social structure (Connor et al., 1998). Females and immature individuals form large social units, called "harem" schools, which are distributed in low and mid-latitudes. Conversely, males have larger ranges, making long migrations between the tropics and high latitudes, even reaching polar waters. Mature males and small groups of sub-adults (bachelor groups) are found in temperate latitudes, which eventually disperse independently into high-latitudes as male breeding bulls (Gaskin, 1970; Whitehead and Weilgart, 2000). Sperm whales mature slowly, with females reaching sexual maturity at about 9 years when roughly 9-m long, and physical maturity when growth ceases at about 11 m (roughly 30 years old) (Best et al., 1984; Rice, 1989). Males grow steadily until their 30 s, starting breeding behavior at their twenties, and reaching physical maturity at about age 50 years when roughly 16-m long (Rice, 1989). Mature males move independently between groups of females for mating and their association with a group can be as brief as several hours or days (Best, 1979; Kahn, 1991; Whitehead, 1993; Whitehead and Arnbom, 1987; Whitehead and Waters, 1990). Seasonality and timing of breeding and whale behavior during the entire year remain largely unknown, with most of the relevant information gathered during summer months (Jochens et al., 2008).

Despite sperm whales being a globally-distributed species, there is some genetic differentiation between populations (Lyrholm and Gyllensten, 1998). A population or "stock" is defined by the Marine Mammal Protection Act (MMPA) as an interbreeding group in a common spatial arrangement, mainly to avoid potential localized depletion due to human effects of mortality and serious injury (Register United States Federal, 2013). The GOM sperm whale population is listed as a separate stock from the U.S Atlantic stock under the MMPA, and the GOM population has genetic, size, and behavior differences with respect to other populations (Engelhaupt et al., 2009; Jaquet, 2006; Jochens et al., 2008). Even decades after commercial whaling was outlawed, the GOM population faces a suite of contemporary threats. The population was exposed to oil from the DWH oil spill and the response activities (i.e. increased vessel and air traffic, vessel strikes,

seismic surveys to detect leaks around the wellhead) (Dias et al., 2017). Additionally, deep-water ambient noise levels in the GOM are among the highest reported in the world's oceans and are persistently high over long periods of time (Wiggins et al., 2016). Animals here are being exposed to high levels of anthropogenic noises generated by geophysical surveys for hydrocarbon deposits, heavy shipping traffic, and large-scale commercial fishery activities (Wright et al., 2007). Through simulation models, Farmer et al. (2018) linked a significant reduction in relative fitness of reproductive females (4% of the stock reaching terminal starvation) to behavioral disturbance associated with hydrocarbon explorations and substantial declines of up to a 25 % of the stock's population.

Sperm whales, like many other species of toothed whales, emit a number of different impulsive sounds when diving (Douglas et al., 2005; Lahiri and Banerjee, 2013; Miller et al., 2008; Watwood et al., 2006; Weilgart and Whitehead, 1990). The most frequently detected sounds are the high amplitude echolocation clicks, with apparent source levels of 245 dB_{pp} re 1µPa at 1m measured from males (Møhl et al., 2003; Zimmer et al., 2005). These unidirectional clicks (roughly 27 dB directivity index; Møhl et al., 2003; Zimmer et al., 2005) are often detected from large distances (range of 10 – 20 km) by an acoustic sensor. These usual clicks contain energy in the 1 – 15 kHz frequency band, and are produced at rates of 0.5 – 2 s throughout 80 % of a foraging dive cycle (Douglas et al., 2005; Watwood et al., 2006; Weilgart and Whitehead, 1990). Usual clicks are interrupted by faster click trains, known as creaks or buzzes, with rates less than 0.2 s (Miller et al., 2004). Sperm whales also produce clicks associated with communication, including click trains called codas and slow clicks produced by males (Jaquet et al., 2001; Madsen et al., 2002; Mullins et al., 1988; Oliveira et al., 2013; Weilgart and Whitehead, 1988, 1990).

Sperm whales mostly feed at mid-sea and deep waters along the continental slopes for squid and fish (Santos et al., 2001; Smith and Whitehead, 2000), where passive acoustic monitoring (PAM) can be useful to observe whale behavior in deep depths and remote habitats. The frequent use of usual clicks by sperm whales makes these signals useful for detecting sperm whale presence. Prior monitoring studies of sperm whales in the GOM have made use of satellite data (Jochens et al., 2008) and acoustic data, including a towed array (Mellinger, 2002), PAM methods (Merkens, 2013), and D-tag data (Jochens et al., 2008).

In this study, long-term passive acoustic data collected using near-seafloor High-frequency Acoustic Recording Packages (HARPs) at three sites along the continental slope of the GOM regions were used to detect the presence of sperm whales for over seven years between 2010 and 2017. The longest-term time series to date is provided documenting the presence of sperm whales in oiled and unoiled areas during and after the DWH oil spill. A method to characterize the sex of sperm whale encounters is described based on inter-click-interval (ICI) distributions of click series and the correlation between ICI and acoustic length of sperm whales. This method is

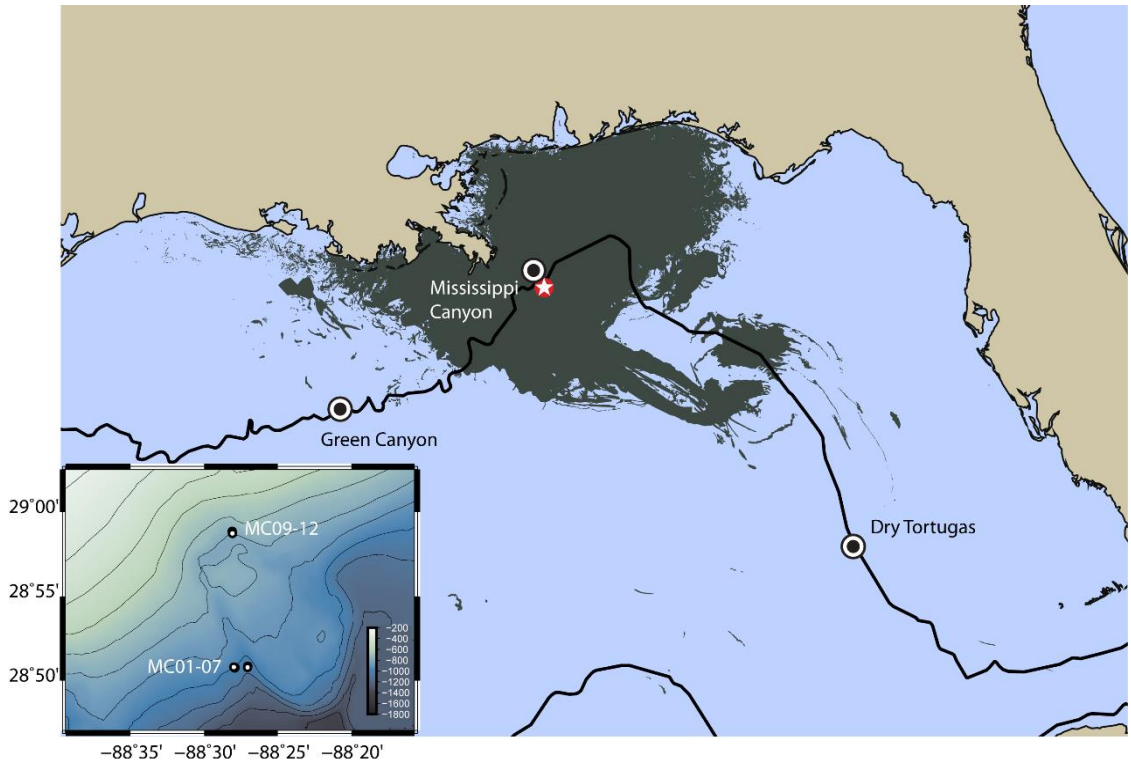


Figure 3.1. Map of deployment locations in the Gulf of Mexico with detections of sperm whales: Green Canyon (GC), Mississippi Canyon (MC), and Dry Tortugas (DT). *Deepwater Horizon* site (red star) and cumulative surface oil during April-August 2010 (dark gray area). The black line denotes the 1000m contour. Surface oil is cumulative NESDIS SAR composite from: http://gomex.erma.noaa.gov. Inset map showed two different locations for monitoring at the MC site. Map generated using GMT (http://gmt.soest.hawaii.edu/projects/gmt).

applied then to the long-term dataset to conduct an investigation of the sperm whale population structure and its spatial and seasonal occupancy of the GOM region.

3.3. Methods

Acoustic data for this study were collected from three deepwater locations in the GOM (**Figure 3.1**) during and following the DWH oil spill (2010-2017). The monitoring locations included a site in Mississippi Canyon (MC) near the DWH wellhead, a western site at Green Canyon (GC) outside of the DWH surface oil footprint, and an eastern site outside of the oil footprint near the Dry Tortugas (DT). At each site, a HARP was deployed and recorded nearly continuously with a sampling frequency of 200 kHz.

Individual sperm whale echolocation clicks were automatically detected using a multi-step approach implemented in MATLAB (Mathworks, Natick, MA). A full description of the algorithm is detailed in **Section 2.3.1.1**. The acoustic encounters were then manually reviewed using the custom graphical user interface *DetEdit*, described in **Section 2.3.2**. Misidentified encounters, from ships or other marine mammals such as beaked whales or delphinids, were removed from the analysis.

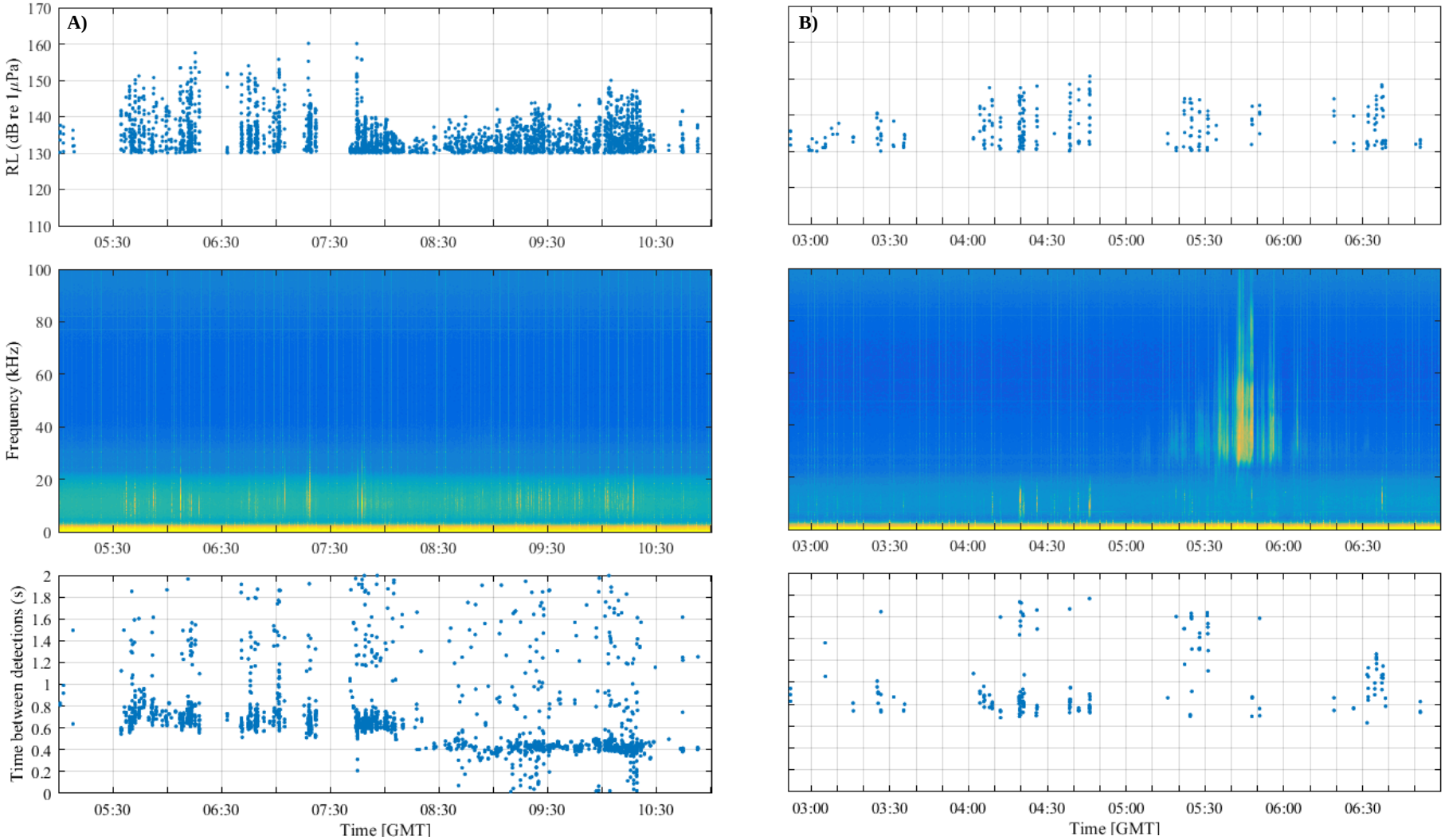


Figure 3.2. Examples of the three distinct ICI patterns (received-level above, long-term spectral average in middle, and ICI below) from two sperm whale encounters in the bottom plots.

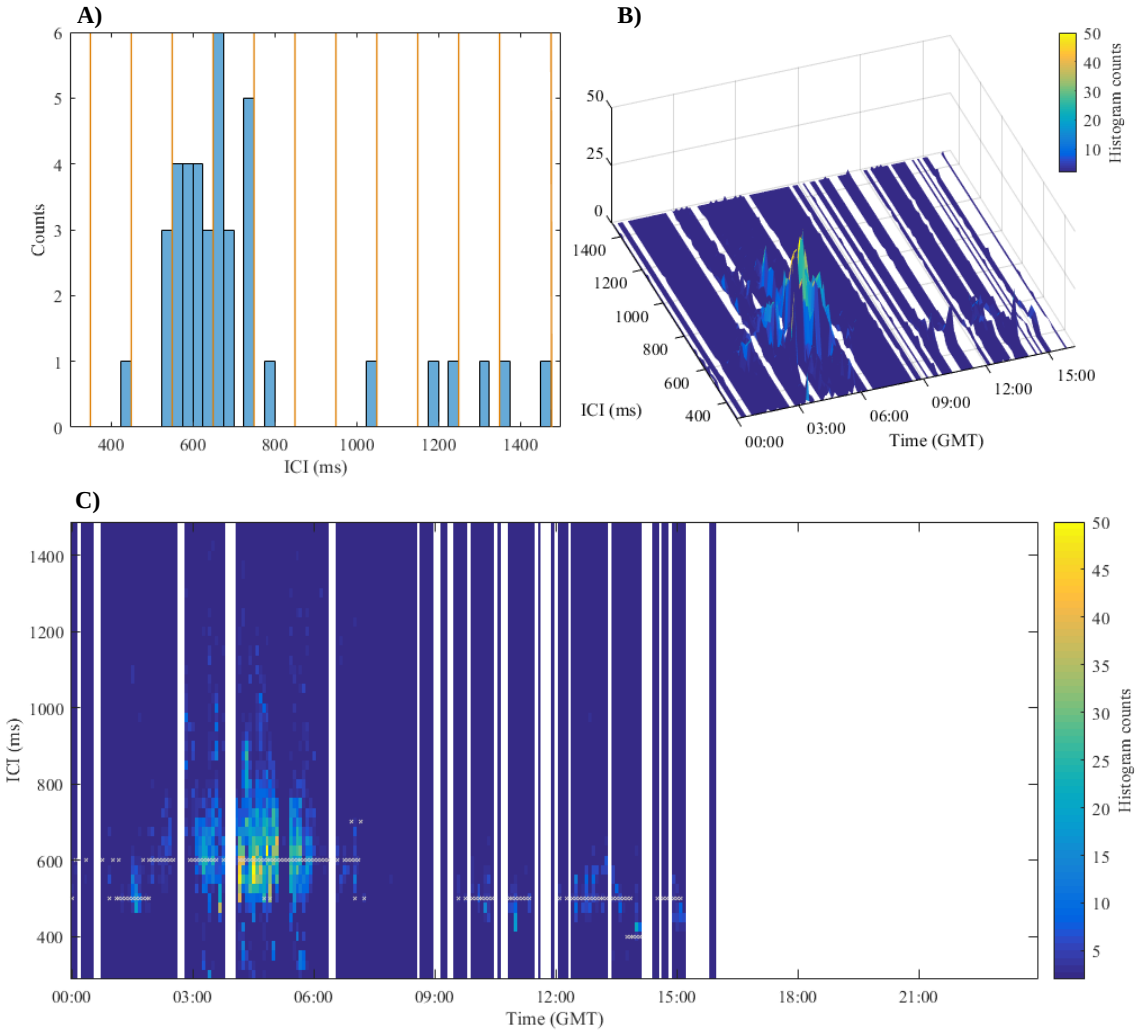


Figure 3.3. Concatenated 2D histograms of ICI distributions (ICIgrams) of detected events during a one-day period. A) A histogram of the ICI distributions with 25 ms bin width over a 5 min interval. B) A 3D representation of the ICIgram plot illustrating the time series of 5 min interval ICI histograms. C) The interface visualization of an ICIgram with successive ICI histograms of one day of PAM data. On the x-axis, histogram counts per 5 min bins are represented by color intensities throughout the span of a day with ICIs on the y-axis. The white points represent the histogram mode during each 5 min interval and indicate over longer periods the dominant ICI bands.

During visual inspection with the *DetEdit* interface, series of clicks of sperm whale encounters with different ICIs were often distinctive (**Figure 3.2**). Encounters included series of clicks with consistent ICIs greater than 0.8 s (**Figure 3.2B**) and conversely ICIs less than 0.8 s (**Figure 3.2A**).

3.3.1. Inter-click-interval classification

A MATLAB-based graphical user interface was developed to determine whether an acoustic encounter of sperm whales was more likely generated from one of the three distinct ICI patterns (**Figure 3.3**). The interface allowed intuitive annotation of detections based solely on ICIs. A distribution of ICIs within a 5 min interval was obtained by accumulating ICIs into a histogram with a discrete bin width of 25 ms. The lower and uppermost bin ranges were defined as 0.3 s and

1.5 s, respectively (**Figure 3.3A**). Values outside this range were ignored to minimize the influence of interleaved trains. For a discrete time span (one day), the ICI histograms were displayed sequentially to generate a surface plot of concatenated histograms, similar to the surface plots referred to as ICIgrams implemented by Miller and Miller, (2018) (**Figure 3.3B**). Comparable to a spectrogram, the ICIgrams allow visualization of the ICI distributions (**Figure 3.3C**).

Reference points indicating the dominant distribution of ICIs in the ICIgrams were used to select the time bins and facilitate the categorization of the selected time following a simple list of keyword shortcut commands. The reference points were obtained by binning the histograms with bin widths of 100 ms, resulting in ten center bins, referred hereafter as ICI band marks (e.g., 400 to 1400 ICI band marks) (**Figure 3.3A**). The ICI band with the highest count was displayed in the ICIgrams as a cross mark reference point. If two ICI bands had the same number of counts, both cross marks were displayed. Following a simple list of keyboard shortcuts and the use of the selection tool, reference points corresponding to a time bin were selected by the analyst and attributed to one of the distinct ICI distributions. Each 5 min-time bin could be attributed to distinct ICI distributions if the distinct patterns were observed on the interface.

Three consistent and distinct ICI distributions were found, representing a regular click interval. The three ICI peak distributions were designated as three different classes, and concatenated ICIgrams of multiple events from these classes are shown in **Figure 3.4**. Class A was characterized by sperm whale encounters with dominant ICI distributions between 0.5 and 0.7 s, class B between 0.6 and 0.8 s, and class C between 0.7 and 0.9 s. Most of the sperm whale encounters were well defined between these three classes. Although in some cases, the dominant ICI distributions were not within the ranges of these three classes. ICI distributions below a distinct range were observed when multiple animals were present, producing overlapping click sequences. ICI distributions above a distinct range were also observed if several clicks within a sequence were not detected. In these cases, it was still obvious which class the dominant ICI distributions fell into. Occasionally, ICI distributions within the range of two of these three classes were observed, causing overlap. In this case, the class was defined according to which range the majority of dominant ICI distributions were seen.

3.3.2. Stable inter-pulse-interval measurements and acoustic body length estimation

The sound produced at the anterior end of the sperm whale's head bounces between reflective air sacs at either end of the spermaceti organ resulting in the inter-pulse-interval (IPI) that composes a single click, being related to twice the travel time for sound along the length of the spermaceti organ (Norris and Harvey, 1972). Thus, IPI is a function of the head length and hence, the total body length of the whale can be estimated (Gordon, 1991; Growcott et al., 2011; Rhinelander and Dawson, 2004). Recordings were analyzed to estimate the stable IPI that composes a single click.

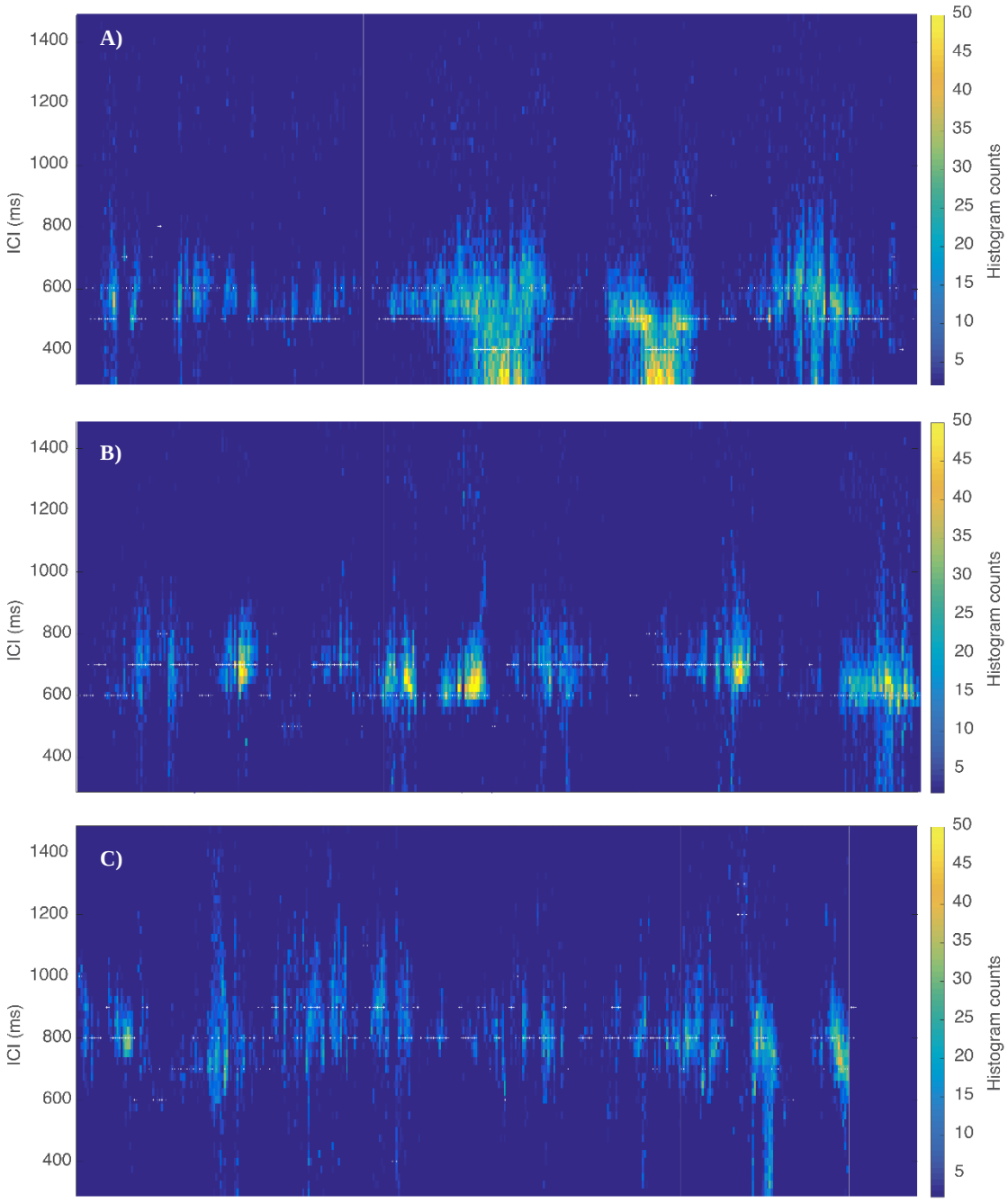


Figure 3.4. Concatenated ICIgram plots of sperm whale encounters representing the different ICI distributions categorized in three different classes, A, B, and C (corresponding plots A, B, and C). The vertical axis indicates the inter-click-interval of the click series and the color bar represents the counts in the histogram for each 25 ms bin.

An effort was made to examine sequences of clicks with ICI distributions between 0.3 and 1.5 s from all deployments of each site, covering the ranges of IPIs reported in literature. The acoustic body length from the sampled clicks was estimated which could be attributed to the known body size of different sexes, and consequently potentially relate the dominant ICI distributions to body size ranges, belonging to maturity stages.

Clicks detected by the automatic detector were manually inspected to measure IPI using a custom software program *Triton* (Wiggins and Hildebrand, 2007), developed in MATLAB (The Mathworks). Long-term spectral averages (LTSAs) were calculated for visual analysis and manual screening for detected acoustic encounters. When echolocation signals were noted in the LTSA, the corresponding spectrogram (1000 point FFT, Hann window, 100 kHz bandwidth) and waveform plots of 10 s length were inspected more closely. To optimize the signal-to-noise ratio, a bandpass filter with a cut-off low end frequency of 5 kHz and high-end frequency of 95 kHz was used. The time between clicks and pulses within the click were determined from the waveform plot using a selection tool within *Triton*. The tool provides the time on the x-axis and amplitude in counts on the y-axis. The ICI and IPI were calculated by subtracting the time points selected in *Triton*. Nearly 500 clicks were examined in the context of over 250 click series. Only click series containing at least five clicks with a consistent ICI, indicating they came from the same whale, were chosen for further examination. Within a click series, only clicks containing at least three pulses with a consistent IPI were used in the final calculations. Since little variation has been observed within IPIs for clicks in a single sequence, unique IPIs were assumed to be from the same whale (Adler-Fenchel, 1980). As noted by several other studies, clicks showing the clear, multi-pulse structure were rare (Adler-Fenchel, 1980; Gordon, 1991; Møhl et al., 1981; Rhinelander and Dawson, 2004) and although there were time delays within the click structure due to off-axis clicks (Zimmer et al., 2005), the whale's IPI was present within each click (Growcott et al., 2011).

Several studies have related IPI estimates to measurements of whale length. Gordon (1991) analyzed IPIs from sperm whales within a nursery group in the Azores and Sri Lanka. The measured whales were relatively small, similar to the population in GOM (Jaquet, 2006).

The acoustic length of the sperm whales was determined using Gordon's equation based on variability in IPIs, IPI trends with time and depth, and vocalizations from whales of known length (Gordon, 1991):

$$Total\ Acoustic\ Length\ = 4.833 + 1.453 \cdot IPI - 0.001 \cdot IPI^2$$

A non-parametric Theil-Sen (Sen, 1968; Theil, 1950) regression, which performs better with outliers, was used to fit a line to the sample points. The slope estimate was used to predict acoustic animal length to the corresponding measured IPI of a click sequence.

3.3.3. Seasonal occupancy

It was also investigated whether there were seasonal trends in the presence of sperm whale detections. Monthly presence was calculated as the percentage of 5-minute intervals in which clicks associated with one of the three classes were detected. Long-term trends in monthly presence were estimated for each class using a Thiel-Sen regression with 5-95% confidence intervals obtained using a bootstrap method. The median slope across 500 pairs of points selected randomly with replacement within each time series was computed 100 times. Seasonal patterns per class were calculated as the monthly presence at each site with respect to the overall mean trend presence of each class.

3.4. Results

3.4.1. Inter-Click-Interval classification

A total of 7367 days of ICIgram plots were visually inspected to define consistent ICI distributions, with 2496 days (7352 h) of detections from site GC, 2558 days (16383 h) from site MC, and 2313 days (2201 h) from site DT. Different distribution of classes was found across sites. Site MC had the highest proportion of class A with 97 %, followed by class B with 2 %, and class C with less than 1%. A similar pattern distribution was seen at site GC, with 95, 4, and 1 % for class A, B, and C, respectively. Contrastingly, at site DT, the proportion of classes varied per time of year, during summer months (April-August) a high proportion of class A with 94 % was observed, followed by class B with 4 %, and class C with 2 %. During the other months, a higher proportion of class B and class C was observed with 20 %, and 8 %, respectively followed by class A with 72 %.

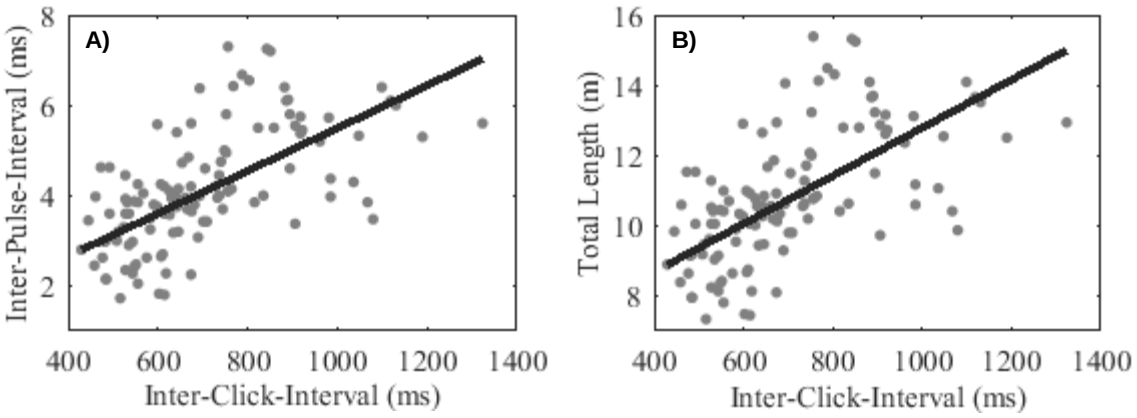


Figure 3.5. Relationship between inter-click-interval and (A) inter-pulse-interval as well as (B) estimated acoustic total length. Each point represents one sampled click sequence with the corresponding mean IPI and estimated animal length. The line fitted (solid line) is a Theil-Sen regression.

3.4.2. Stable Inter-Pulse-Interval measurements and acoustic body length estimation

A total of 116 clicks series met the criteria of having five clicks with a consistent ICI, with 30 clicks series identified from site GC, 37 from site MC, and 54 from site DT. Between all three sites, ICIs ranged between 0.43 s and 1.32 s and IPIs ranged between 1.73 ms and 7.30 ms, which indicate acoustic lengths between 7.34 m and 15.39 m. This is consistent with other sperm whale size ranges reported in the GOM (7.1 – 12.3 m for females and immature sperm whales, Jaquet, 2006; Jochens et al., 2008). A linear correlation between animal length and ICI showed that a 100 ms increase of ICI represented a 1 m increase of total animal length. Based on these results, animals with an acoustic length below 12 m were observed to use mean ICIs below 0.7 s (**Figure 3.5B**). Therefore, class A (with dominant ICIs between 0.5 s and 0.7 s) was hypothesized to correspond to social units (or mixed groups) consisting of adult females and their immature offspring of typical size below 12 m. The size of males in the GOM has been estimated for immature, but not mature individuals (Jochens et al., 2008). As shown in **Figure 3.5B**, multiple animals with an estimated acoustic length above 12 m were sampled, corresponding to a mean ICI above 0.8 s. Therefore, class C (with dominant ICIs of 0.7 s to 0.9 s) was hypothesized to correspond to adult males. These encounters were mostly produced by single animals, likely large

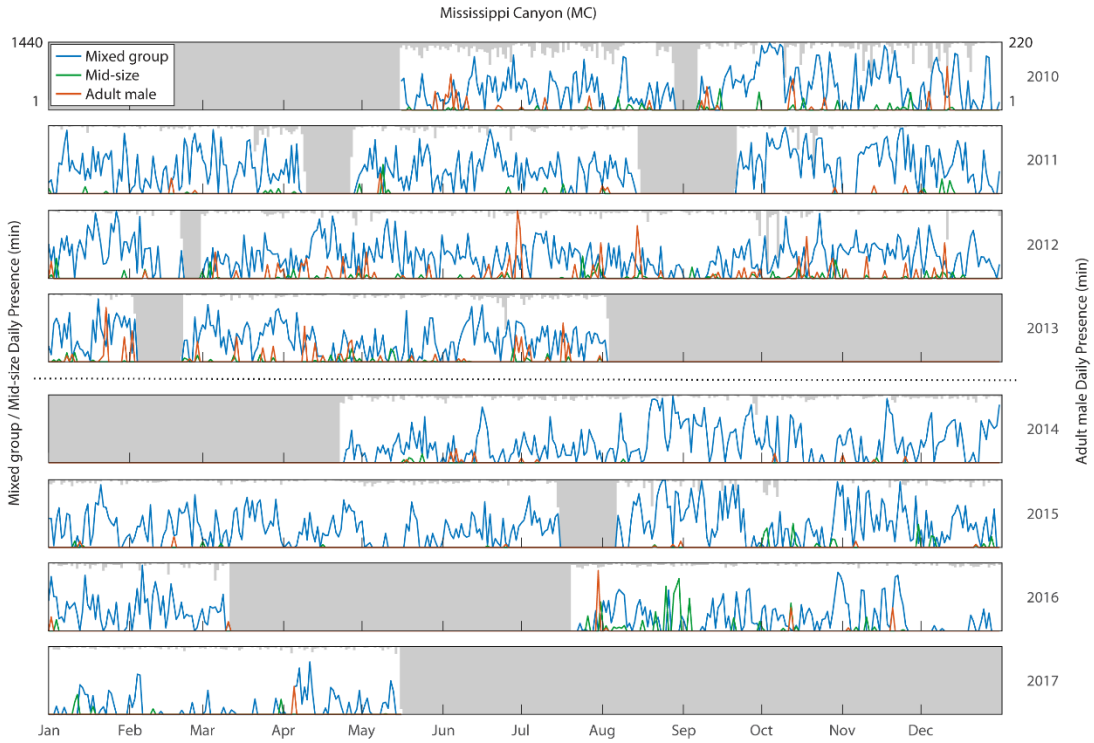


Figure 3.6. Daily presence of sperm whale detections at site MC between 2010 and 2017. The blue line indicates presence of sperm whale encounters categorized as class A (mixed group), green line as class B (mid-size), and red line as class C (adult males). The gray area shows times of no effort data. The dashed line indicates the time when the acoustic recorder was moved from the southwest of the seamount at mean depth of 980 m to the northern site at mean depth of 800 m.

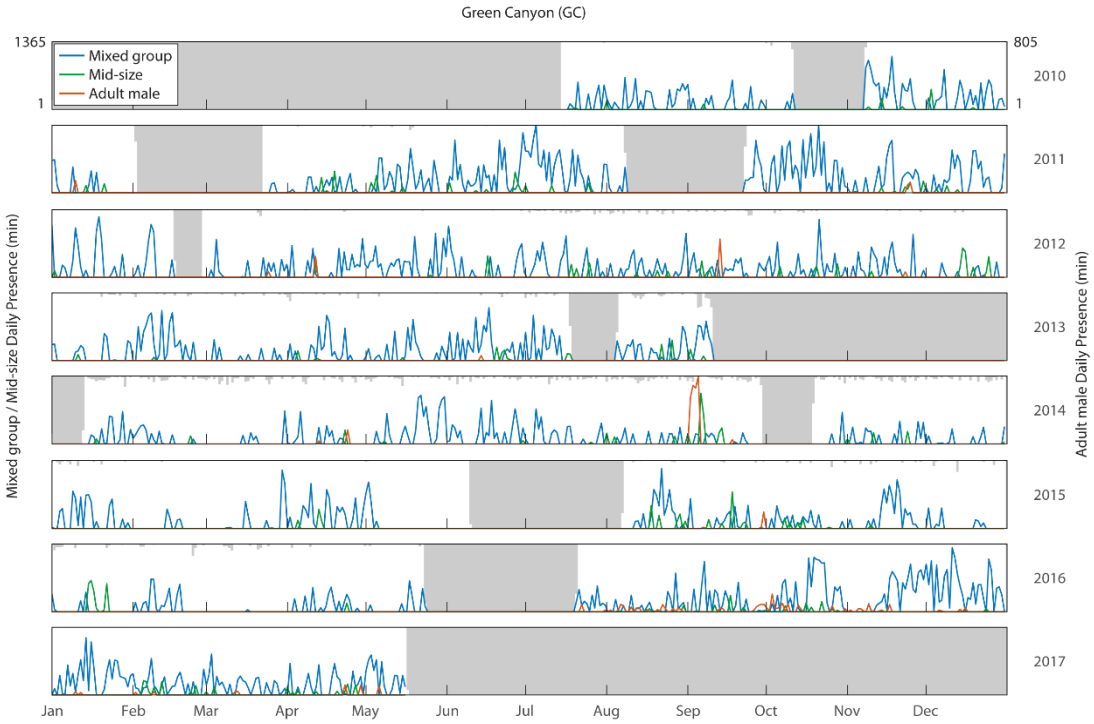


Figure 3.7. Daily presence of sperm whale detections at site GC between 2010 and 2017. The blue line indicates presence of sperm whale encounters categorized as class A (mixed group), green line as class B (mid-size), and red line as class C (adult males). The gray area shows times of no effort data.

solitary animals, which is typical behavior for breeding bulls. Although class B was easily distinguished in the ICIgrams, the length ranges for class B encounters overlapped with those of the other two classes. Hence, it was not possible to define a distinct maturity class for class B based on the relationship of ICI and acoustic length. For purposes of comparison, class B will be referred to as the mid-size class (larger than A and smaller than C).

3.4.3. Temporal occurrence of sperm whale echolocation clicks

Sperm whales were detected often at all three GOM sites during the seven years of monitoring. The three distinct ICI classes were also detected across all three sites with notable differences in temporary use of the three sites (**Figure 3.6**,**Figure 3.7**,**Figure 3.8**). At the level of daily presence, class A associated with mixed groups, were present most frequently (64 % of recording days) at site MC throughout the year. At site GC, there were slightly fewer detections of class A encounters, with clicks present on 42 % of recording days. Interestingly, class A encounters were predominantly present in summer months (May-August) at site DT, with notably fewer detections than the other sites (clicks present on just 14 % of recording days). Class B encounters, referred as mid-size class, were present at all sites throughout the year with notably fewer detections than class A. Class B encounters were present in 6 %, 8 %, and 4 % of recording days at sites GC, MC, and DT, respectively. Class C encounters, associated with adult males, had the lowest presence at all sites, with clicks present sporadically all throughout the year in 3 %, 6 %, and 3 % of recording

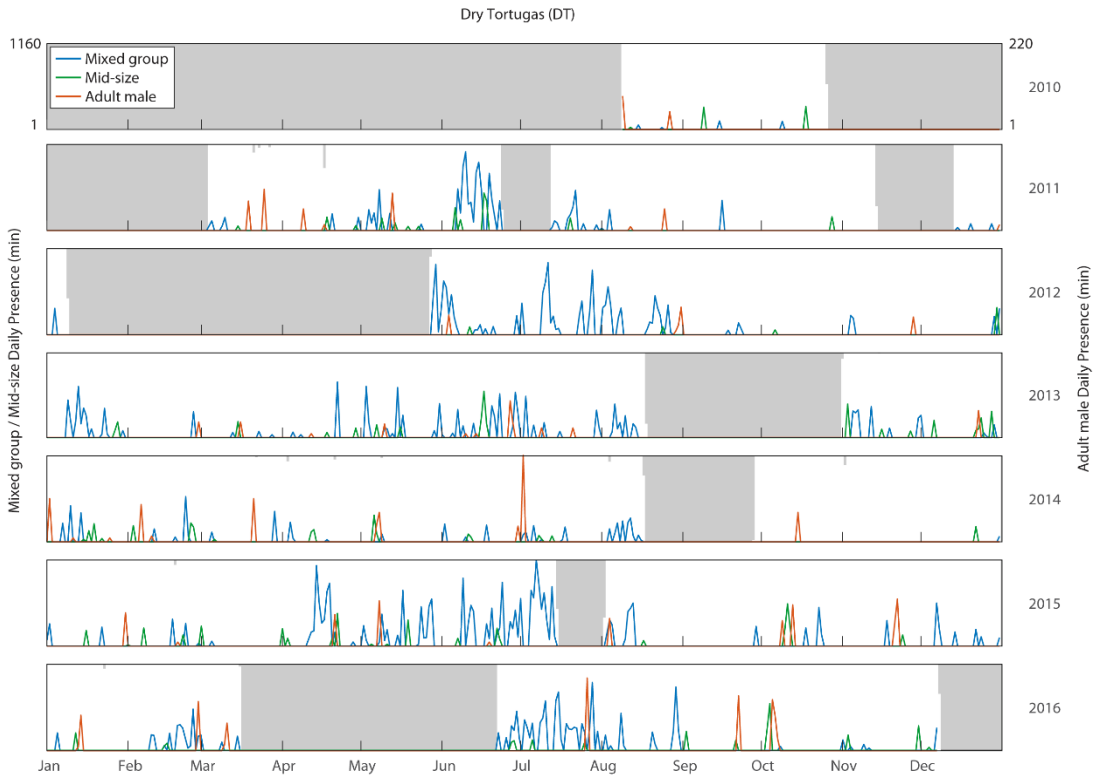


Figure 3.8. Daily presence of sperm whale detections at site DT between 2010 and 2016. The blue line indicates presence of sperm whale encounters categorized as class A (mixed group), green line as class B (mid-size), and red line as class C (adult males). The gray area shows times of no effort data.

days at GC, MC, and DT, respectively. There was a considerable decrease in detections of adult males at site MC after 2013, likely due to a change in the location of the acoustic recorder (**Figure 3.1**). During 2010 – 2013, the hydrophone was deployed at a mean depth of 980 m southwest of a seamount. However, starting in 2014, the hydrophone was deployed 15 km north of the previous location, located north of the seamount at a depth of 800 m.

Basic statistical measures of the seasonal events for the seven years of data per site and class were evaluated with respect to the overall mean trend of each class (**Figure 3.9**). The presence of class A (mixed groups) was found to slightly decrease in late winter (February-March) at site GC and in spring (April-May) at site MC. However, at site DT, their presence increased during spring and summer months (April-August). The presence of class B encounters was observed to increase during summer and fall (July-October) at the northern sites, with a simultaneous decrease at the eastern site. Similarly, at site GC, class C (adult males) was detected more often during later summer and fall (July-November). Although the peak presence of class C encounters occurred earlier in summer (June-July) at site MC, the eastern site did not show any evident variability across months.

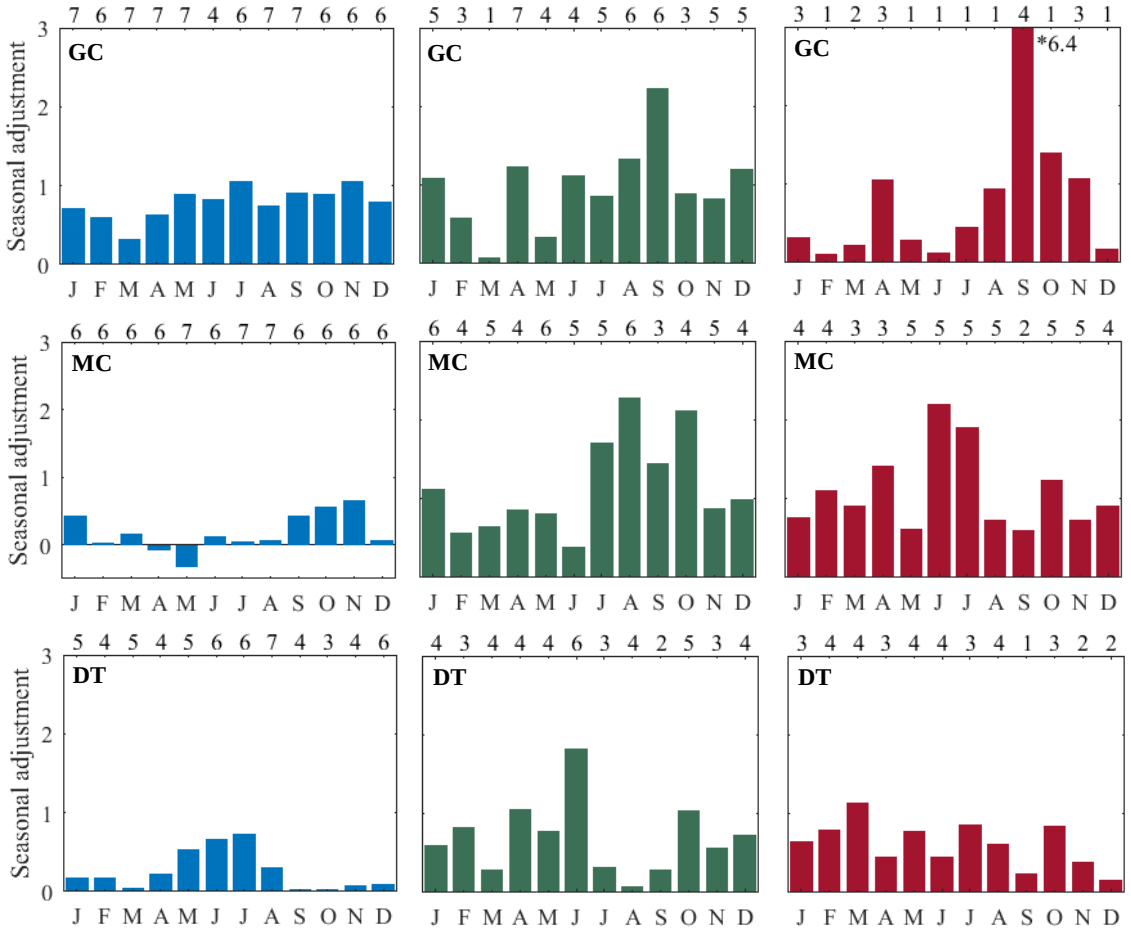


Figure 3.9. Seasonal patterns of sperm whales as a function of month for each site and sex-class. Blue plots (left column) indicate patterns for Class A (mixed groups), green plots (center column) indicate Class B (mid-size), and red plots (right column) indicate Class C (adult males) for the three monitoring sites, GC, MC, and DT. The vertical axis indicates the factor by which seasonal presence varies relative to mean presence of the class for all sites. Higher values indicate stronger seasonality. The numbers on the top of the plots represent the number of years with presence per each month.

3.5. Discussion

This study provides the longest time series of sperm whale occurrence using passive acoustic monitoring to date. The presence of sperm whale regular clicks was detected along the continental slope at three sites in the GOM, in oiled and unoiled areas, during and after the DWH oil spill. The results indicate spatial and seasonal variation in the population structure of the GOM sperm whale population.

A subsample of sperm whale detections (116 click series) was characterized and estimates of size were given based on measurements of the IPI within a click series. Although it has been reported that IPI can vary with function of depth (Madsen, 2002), Bøttcher et al. (2018) concluded that this variation is only on the order of 0.2 ms and a size error estimate of about 0.3 m. Considering this variability, sizes in our study were consistent to those reported for the GOM region (Jaquet, 2006;

Jochens et al., 2008), although sizes for mature males have hardly been reported for this area (Collum and Fritts, 1985; Miller, 2004). Contrastingly, numerous animals longer than 12 m and up to 15 m were detected at the study sites. Solitary animals seen from the air have been estimated to be as large as 14 – 15 m (Collum and Fritts, 1985), thus giving support for the presence of breeding mature males in the GOM.

The method developed and implemented as an interface to manually categorize sperm whale events into possible sexes, based on ICI distributions, allows for expeditious processing of over seven thousand days of sperm whale events. The method worked well in allowing the consistent categorization of events, even with multiple animals clicking or clicks being missed in a click series. Three consistent and distinct ICI distributions, categorized as class A, B, and C, were found and related to the linear correlation between the ICI and animal length. Previously, Goold and Jones, (1995) noted that adult males and females in the waters around the Azore Islands had different mean ICIs of 0.85 s and 0.51 s, respectively. The size estimates of the Azores sperm whales (Gordon, 1991) are similar to those found in the GOM population (Jaquet, 2006). The ICI distribution that we observed is in close agreement with the values reported in Gordon's study, demonstrating that the different, consistent ICIs could be a good acoustic indicator to determine the sex-class of the detected encounters.

Based on the linear correlation of ICI and animal size, events with ICI distributions between 0.5 – 0.7 s (class A) corresponded to animals of sizes below 12 m. Therefore, these detected encounters were hypothesized to be produced by small whales (7 to 12m) corresponding to mixed groups of adult females and their immature offspring. These sizes are consistent with those reported by Jaquet, (2006) and Jochens et al. (2008), where adult females are on average 1.5 – 2.0 m smaller than the global mean. Distinctively, consistent ICI distribution between 0.7 – 0.9 s (class C) corresponded to animals of sizes above 12 m. Based on one tagged adult male of 12.4 m in the GOM (Miller, 2004) and the solitary animals seen from the air of 14 – 15 m length (Collum and Fritts, 1985), it was assumed that class C is large whales (12 – 15 m) corresponding to adult males. Furthermore, the large males (14 – 16 m) observed in the data are similar to those reported in the Atlantic (Miller, 2004; Santos et al., 1999). These measurements are consistent with the hypothesis that males from the Atlantic population move in and out of the GOM (Engelhaupt et al., 2009). Although measurements are obtained from a correlation based from a manual analysis of a subset of data and size error estimates are to be expected.

Even though consistent ICI distributions between 0.6 – 0.8 s (class B) were well differentiated when visually inspecting ICIgram plots, corresponding animal sizes to these ICI ranges were not distinct enough to determine sexual maturity. Animal sizes in this class correspond to both big adult females and juveniles or sub-adult males. Given their distinct separation from mixed groups, the mid-size class is suggestive of the presence of sub-adult male groups, which are often found

in temporary aggregations of smaller “bachelors” schools by themselves. Future analysis, like the approach implemented by Caruso et al., (2015), using cepstrum analysis to automatically measure the stable IPIs of detected clicks with the three distinct ICI distributions could provide more accurate size ranges within each class.

Home ranges and core areas of sperm whales in the GOM were identified using satellite tracking locations, although this ranges were identified from the movements of a limited sample size of 52 tagged sperm whales of adult females and juveniles (Jochens et al., 2008). The home ranges compromised nearly the entire Gulf in waters deeper than 500 m and a small portion of the southeast North Atlantic coast. Core areas were located in the northern slope range of the Gulf, this being the boundaries of the female’s home range. Core areas of females were located in the north western and central Gulf, near the study site MC. However, site GC, was located between the two identified core areas, along the 1000-m contour, and was frequently used by sperm whales. Similar to these observations, the presence of mixed groups in the present study indicates a high usage of the northern sites year-round; site MC being the most utilized. It was reported that movements of sperm whales in the northern GOM were not migratory, but were irregular likely linked to changes in food availability (Jochens et al., 2008). However, the presence of the mixed group class was observed to increase, almost doubling the overall mean trend of this class at site DT during spring and summer months for almost all the monitored years, indicating, a migration pattern. A slight decrease was observed in late winter at site GC and in spring at site MC. It is believed that the significant seasonal pattern at site DT was likely driven by a portion of whales that left the northern sites and migrated to the eastern site. The reasons for this distinct migration pattern could be linked to oceanographic conditions associated with prey availability or breeding, where females are moving closer to where males possibly move in an out of the Gulf. However, no peak presence of adult males was observed at site DT. Instead, males were sporadically present throughout the year, making the reason for this migratory pattern more complicated. The presence of adult males was also observed sporadically throughout the year at site MC, the most frequented of all sites. More than doubling of the adult male class was observed during June and July at site MC and during fall at site GC. Site GC is located along the transit corridor between both female core areas in the northern Gulf, and the behavior exhibited by the males at this site could be linked to transit behavior.

One tagged male, likely not mature (Jochens et al., 2008), moved out of the Gulf and used the most southern area of the Gulf (the northwest Cuban coast) to move into the Atlantic ocean. This could explain why the seasonal pattern observed at site GC is not observed at site DT although the site was frequented year-round by the adult male class. These results suggest that like in the South Pacific and in the Gulf of California, the breeding season is extensive, encompassing most

months of the year where males rove between groups of females staying only a few hours or days with each group (Coakes, 2004; Jaquet and Gendron, 2009; Whitehead, 1990).

Adult males were detected at the southeast side of the seamount at site MC at a mean depth of 980 m. However, when the acoustic recorder was moved to the northern side of the seamount at a mean depth of 800 m in 2014, the presence of adult males was reduced drastically. This indicates that males mostly make use of deep waters. Contrastingly, the mixed groups did use both sides of the seamount, with a higher presence at the deeper site.

Although the mid-size class has not been attributed to a specific sexual maturity class, some patterns of presence and movement were observed. The presence of the mid-size class was higher compared to the adult male class at all sites, with site MC being the most frequented site. Although presence was sporadic year-round at all sites, an increase during late summer and fall months of more than double the mean trend of this class was observed at site MC, corresponding to a decrease at site DT. Site GC only shows a peak in presence in September. This could indicate that a portion of whales move in late summer from the southern site to the northern Gulf, or the core area at MC.

3.6. Conclusions

A method to categorize sperm whale events into possible sex classes based on ICI distributions is presented. It is also shown how the correlation of ICI and acoustic length allows for the categorization of sexes from sperm whale events using seven years of acoustic data. This method provides a means to efficiently characterize the population structure of sperm whales. Application of this method has revealed new insight into the understanding of breeding, spatial, and seasonal variability of the GOM sperm whale population. The capability to acquire data for an extended period of time with passive acoustic monitoring provided advantages and an increase in our knowledge about this deep-diving cetacean species. The results of this study are a precursor to an acoustic estimate of population density of sperm whales in the Gulf of Mexico, providing additional information about the population structure which facilitates the estimation of cue rates and other multipliers required for density estimation.

3.7. Acknowledgments

This research was made possible by grants from The Gulf of Mexico Research Initiative through its consortium The Center for the Integrated Modeling and Analysis of the Gulf Ecosystem (C-IMAGE). Funding for acoustic data collection and analysis was also provided by the Natural Resource Damage Assessment partners (20105138), the US Marine Mammal Commission (20104755/E4061753), the Southeast Fisheries Science Center under the Cooperative Institute for Marine Ecosystems and Climate (NA10OAR4320156) with support through Interagency Agreement #M11PG00041 between the Bureau of Offshore Energy Management, Environmental Studies Program and the National Marine Fisheries Service, Southeast Fisheries Science Center. The data used for this study are archived by the Gulf of Mexico Research Initiative at https://data.gulfresearchinitiative.org/data/R4.x267.180:0011 maintained by the Gulf Research Initiative Information and Data Cooperative. I thank Steve Murawski, Sheryl Gilbert, Melissa Soldevilla and Lance Garrison for their help with this work.

3.8. References

Adler-Fenchel, H. S. (**1980**). "Acoustically Derived Estimate of the Size Distribution for a Sample of Sperm Whales (Physeter catodon) in the Western North Atlantic," Can. J. Fish. Aquat. Sci., doi: 10.1139/f80-283.

Best, P. B. (**1979**). "Social Organization in Sperm Whales, Physeter macrocephalus," In H. E. Winn and B. L. Olla (Eds.), Behav. Mar. Anim. Vol. 3, Plenum Press, New York, pp. 227–290.

Best, P. B., Canham, P. A. S., and Macleod, N. (**1984**). "Patterns of reproduction in sperm whales, Physeter macrocephalus," Reports Int. Whal. Commision, doi: 10.1021/acs.jpcc.5b11294.

Bøttcher, A., Gero, S., Beedholm, K., Whitehead, H., and Madsen, P. T. (**2018**). "Variability of the inter-pulse interval in sperm whale clicks with implications for size estimation and individual identification," J. Acoust. Soc. Am., **144**, 365–374.

Caruso, F., Sciacca, V., Bellia, G., Domenico, E. De, Larosa, G., Papale, E., Pellegrino, C., et al. (**2015**). "Size Distribution of Sperm Whales Acoustically Identified during Long Term Deep-Sea Monitoring in the Ionian Sea Size Distribution of Sperm Whales Acoustically Identified during Long Term Deep-Sea Monitoring in the Ionian Sea," PLoS One, **10**, 1–16.

Coakes, A. (**2004**). "Social structure and mating system of sperm whales off northern Chile," Can. J. Zool., **82**, 1360–1369.

Collum, L. A., and Fritts, T. H. (**1985**). "Sperm Whales (Physeter catodon) in the Gulf of Mexico," Southwest. Nat., **30**, 101–104.

Connor, R. C., Mann, J., Tyack, P. I., and Whitehead, H. (**1998**). "Social evolution in toothed whales," Trends Ecol. Evol., **13**, 228–232.

Dias, L. A., Litz, J., Garrison, L., Martinez, A., Barry, K., and Speakman, T. (**2017**). "Exposure of cetaceans to petroleum products following the Deepwater Horizon oil spill in the Gulf of Mexico," Endanger. Species Res., **33**, 119–125.

Douglas, L. A., Dawson, S. M., and Jaquet, N. (**2005**). "Click rates and silences of sperm whales at Kaikoura, New Zealand," J. Acoust. Soc. Am., doi: 10.1121/1.1937283.

Engelhaupt, D., Rus Hoelzel, A., Nicholson, C., Frantzis, A., Mesnick, S., Gero, S., Whitehead, H., et al. (**2009**). "Female philopatry in coastal basins and male dispersion across the North Atlantic in a highly mobile marine species, the sperm whale (Physeter macrocephalus)," Mol. Ecol., **18**, 4193–4205.

Farmer, N. A., Baker, K., Zeddies, D. G., Denes, S. L., Noren, D. P., Garrison, L. P., Machernis, A., et al. (**2018**). "Population consequences of disturbance by offshore oil and gas activity

for endangered sperm whales (Physeter macrocephalus)," Biol. Conserv., **227**, 189–204.

Gaskin, D. E. (**1970**). "Composition of schools of sperm whales physeter catodon linn East of New Zealand," New Zeal. J. Mar. Freshw. Res., **4**, 456–471.

Goold, J. C., and Jones, S. E. (**1995**). "Time and frequency domain characteristics of sperm whale clicks," J. Acoust. Soc. Am., **98**, 1279–1291.

Gordon, J. C. D. (**1991**). "Evaluation of a method for determining the length of sperm whales (Physeter catodon) from their vocalizations," J. Zool., **224**, 301–314.

Growcott, A., Miller, B., Sirguey, P., Slooten, E., and Dawson, S. (**2011**). "Measuring body length of male sperm whales from their clicks: The relationship between inter-pulse intervals and photogrammetrically measured lengths," J. Acoust. Soc. Am., **130**, 568–573.

Jaquet, N. (**2006**). "A simple photogrammetric technique to measure sperm whales at sea," Mar. Mammal Sci., **22**, 862–879.

Jaquet, N., Dawson, S., and Douglas, L. (**2001**). "Vocal behavior of male sperm whales: Why do they click?," J. Acoust. Soc. Am., doi: 10.1121/1.1360718.

Jaquet, N., and Gendron, D. (**2009**). "The social organization of sperm whales in the Gulf of California and comparisons with other populations," J. Mar. Biol. Assoc. United Kingdom, **89**, 975.

Jochens, A., Jochens, A., Biggs, D., Biggs, D., Benoit-Bird, K., Benoit-Bird, K., Engelhaupt, D., et al. (**2008**). "Sperm whale seismic study in the Gulf of Mexico Synthesis Report," Mms, **96**, 3268.

Kahn, B. (**1991**). *The population biology and social organization of sperm whales (Physeter macrocephalus) off the Seychelles: indications of recent exploitation* Dalhousie University, Halifax, N.S.

Lahiri, K., and Banerjee, S. (**2013**). "India's Health-care Sector under GATS: Inquiry into Backward and Forward Linkages," Foreign Trade Rev., **48**, 285–357.

Lyrholm, T., and Gyllensten, U. (**1998**). "Global matrilineal population structure in sperm whales as indicated by mitochondrial DNA sequences," Proc. R. Soc. B Biol. Sci., **265**, 1679–1684.

Madsen, P. T. (**2002**). *Morphology of the sperm whale head: A review and some new findings* University of Aarhus.

Madsen, P. T., Wahlberg, M., and Møhl, B. (**2002**). "Male sperm whale (Physeter macrocephalus) acoustics in a high-latitude habitat: Implications for echolocation and communication," Behav. Ecol. Sociobiol., doi: 10.1007/s00265-002-0548-1.

Mellinger, D. (**2002**). "Passive acoustic monitoring of sperm whales in the Gulf of Mexico, with a model of acoustic detection distance," … -first Annu. Gulf Mex. …, Retrieved from ftp://heceta.pmel.noaa.gov/newport/mellinger/papers/MellingerEtAl03-SpermMonitoring+DetDistance.pdf.

Merkens, K. (**2013**). *Deep-Diving Cetaceans of the Gulf of Mexico: Acoustic Ecology and Response to Natural and Anthropogenic Forces Including the Deepwater Horizon Oil Spill* University of California, San Diego, 146 pages.

Miller, B. S., and Miller, E. J. (**2018**). "The seasonal occupancy and diel behaviour of Antarctic sperm whales revealed by acoustic monitoring," Sci. Rep., **8**, 5429.

Miller, P. J. O. (**2004**). "Swimming gaits, passive drag and buoyancy of diving sperm whales Physeter macrocephalus," J. Exp. Biol., **207**, 1953–1967.

Miller, P. J. O., Aoki, K., Rendell, L. E., and Amano, M. (**2008**). "Stereotypical resting behavior of the sperm whale," Curr. Biol.,.

Miller, P. J. O., Johnson, M. P., and Tyack, P. L. (**2004**). "Sperm whale behaviour indicates the use of echolocation click buzzes 'creaks' in prey capture," Proc. Biol. Sci., **271**, 2239–47.

Møhl, B., Larsen, E., and Amundin, M. (**1981**). "Sperm whale size determination: Outlines of an acoustic approach," FAO Fish. Ser., **5**, 327–332.

Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., and Lund, A. (**2003**). "The monopulsed nature of sperm whale clicks," J. Acoust. Soc. Am., **114**, 1143.

Mullins, J., Whitehead, H., and Weilgart, L. S. (**1988**). "Behavior and vocalizations of two single sperm whales, Physeter macrocephalus, off Nova Scotia," Can. J. Fish. Aq. Sci., **45**, 1736–1743.

National Academic of Sciences Engineering and Medicine (**2017**). *Approaches to Understanding the Cumulative Effects of Stressors on Marine Mammals*, National Academies Press, Washington, D.C.

Norris, K. S., and Harvey, G. W. (**1972**). "A theory for the function of the spermaceti organ of the sperm whale (Physeter macrocephalus)," In S. R. Galler, K. Schmidt-Koening, G. J. Jacobs, and R. E. Belleville (Eds.), Anim. Orientat. Naveg., NASA, Washington D.C., Special Pu., pp. 397–419.

NRDA (**2016**). "Deepwater Horizon Oil Spill Final Programmatic Damage Assessment and Restoration Plan and Final Programmatic Environmental Impact Statement.,"

Oliveira, C., Wahlberg, M., Johnson, M., Miller, P. J. O., and Madsen, P. T. (**2013**). "The function of male sperm whale slow clicks in a high latitude habitat: Communication, echolocation,

or prey debilitation?," J. Acoust. Soc. Am., **133**, 3135–3144.

Register United States Federal (**2013**). *Endangered and threatened wildlife*, (National Marine Fisheries Service (NMFS), Ed.) Commerce National Oceanic and Atmospheric Administration, United States of America, Notice of., 68032-68037 pages.

Rhinelander, M. Q., and Dawson, S. M. (**2004**). "Measuring sperm whales from their clicks: Stability of interpulse intervals and validation that they indicate whale length," J. Acoust. Soc. Am., **115**, 1826–1831.

Rice, D. W. (**1989**). "Sperm whale Physeter macrocephalus Linnaeus," In S. H. Ridgeway and R. Harrison (Eds.), Handb. Mar. Mamm., Academic Press, San Diego, pp. 177–233.

Santos, M. B., Clarke, M. R., and Pierce, G. J. (**2001**). "Assessing the importance of cephalopods in the diets of marine mammals and other top predators: Problems and solutions," Fish. Res., doi: 10.1016/S0165-7836(01)00236-3.

Santos, M. B., Pierce, G. J., Boyle, P. R., and Reid, R. J. (**1999**). "Stomach contents of sperm whales Physter macropcephalus stranded in the North Sea 1990-1996," Mar. Ecol. Prog. Ser., **183**, 281–294.

Sen, P. K. (**1968**). "Estimates of the Regression Coefficient Based on Kendall's Tau," J. Am. Stat. Assoc., **63**, 1379–1389.

Smith, S. C., and Whitehead, H. (**2000**). "The diet of galapagos sperm whales Physeter macrocephalus as inicated by fecal sample analysis," Mar. Mammal Sci., **16**, 315–325.

Theil, H. (**1950**). "A Rank-Invariant Method of Linear and Polynomial Regression Analysis," Ned. Akad. Wetenchappen Ser. A, **53**, 386–392.

Watwood, S. L., Miller, P. J. O., Johnson, M., Madsen, P. T., and Tyack, P. L. (**2006**). "Deep-diving foraging behaviour of sperm whales (Physeter macrocephalus)," J. Anim. Ecol., **75**, 814–825.

Weilgart, L. S., and Whitehead, H. (**1988**). "Distinctive vocalizations from mature male sperm whales (Physeter macrocephalus)," Can. J. Zool., doi: 10.1139/z88-282.

Weilgart, L., and Whitehead, H. (**1990**). "Click rates from sperm whales," J. Acoust. Soc. Am., doi: 10.1121/1.399376.

Whitehead, H. (**1990**). "Rules for roving males," J. Theor. Biol., **145**, 355–358.

Whitehead, H. (**1993**). "The Behavior of Mature Male Sperm Whales on the Galapagos-Islands Breeding Grounds," Can. J. Zool. Can. Zool., **71**, 689–699.

Whitehead, H., and Arnbom, T. (**1987**). "Social organization of sperm whales off the Galapagos

Islands, February–April 1985," Can. J. Zool., doi: 10.1139/z87-145.

Whitehead, H., and Waters, S. (**1990**). "Social organisation and population structure of sperm whales off the Galapagos Islands, Ecuador (1985 and 1987)," Reports Int. Whal. Comm. Spec. Issue, **12**, 249–257.

Whitehead, H., and Weilgart, L. (**2000**). "The Sperm whale: Social females and roving males," In J. Mann, R. C. Connor, P. L. Tyack, and H. Whitehead (Eds.), Cetacean Soc. F. Stud. Dolphins Whales., University of Chicago Press, Chicago, pp. 154–172.

Wiggins, S. M., Hall, J. M., Thayre, B. J., and Hildebrand, J. A. (**2016**). "Gulf of Mexico low-frequency ocean soundscape impacted by airguns," J. Acoust. Soc. Am., **140**, 176–183.

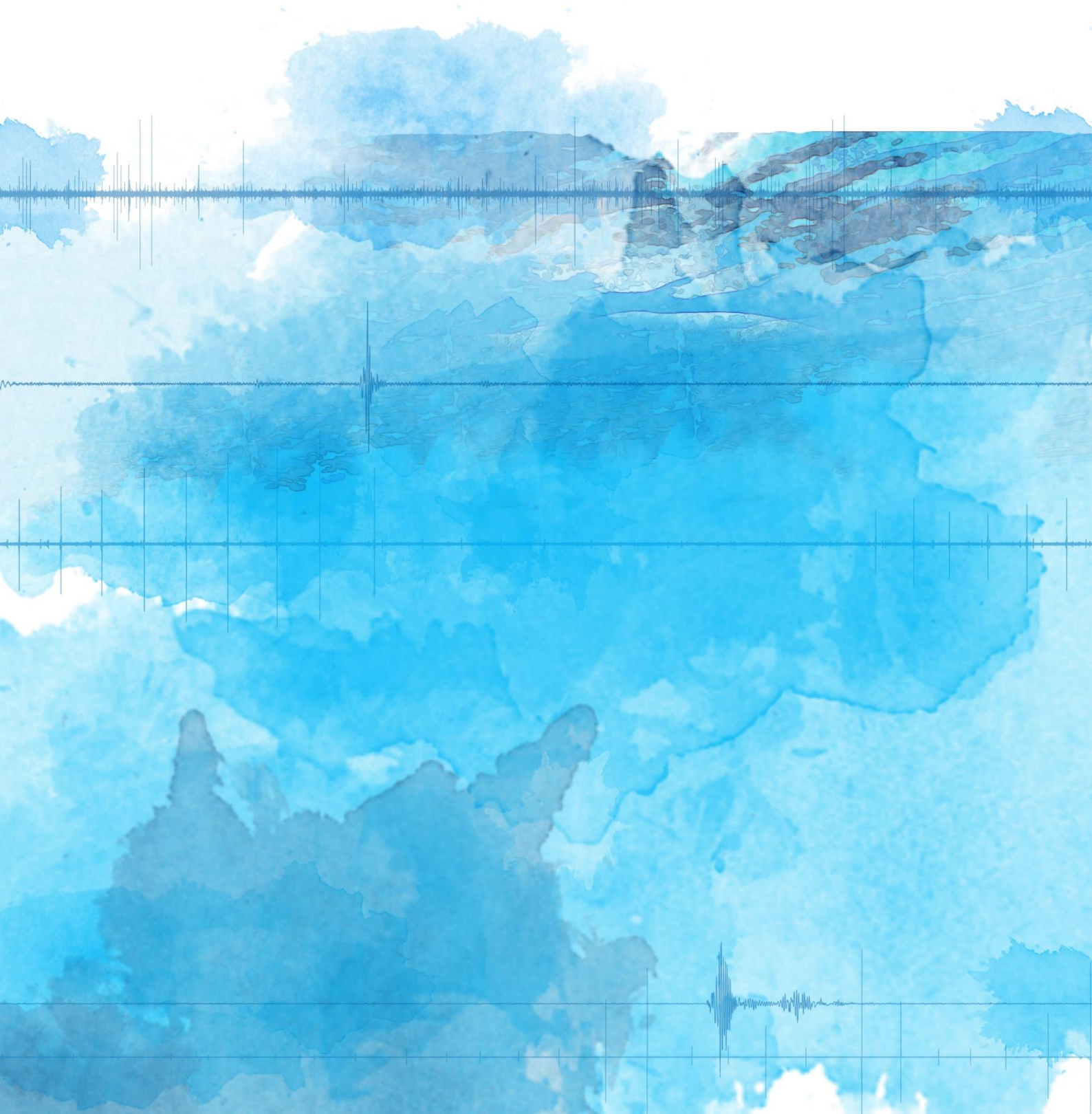
Wiggins, S. M., and Hildebrand, J. A. (**2007**). "High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring," Int. Symp. Underw. Technol. 2007 Int. Work. Sci. Use Submar. Cables Relat. Technol. 2007, Institute of Electrical and Electrongics Engineers, Tokyo, Japan, 551–557.

Wright, A. J., Soto, N. A., Baldwin, A. L., Bateson, M., Beale, C. M., and Clark, C. (**2007**). "Anthropogenic Noise as a Stressor in Animals: A Multidisciplinary Perspective," Int. J. Comp. Psychol., Retrieved from https://escholarship.org/uc/item/46m4q10x.

Zimmer, W. M. X., Tyack, P. L., Johnson, M. P., and Madsen, P. T. (**2005**). "Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis," J. Acoust. Soc. Am., **117**, 1473.

Chapter 4

Sperm whale density trends in the Gulf of Mexico over seven years of passive acoustic monitoring



4.1. Abstract

The Gulf of Mexico's underwater soundscape is intensely impacted by human activities, and an important habitat for globally endangered sperm whales. The Gulf of Mexico is home to heavily-trafficked shipping ports, significant commercial fishery activity, and hydrocarbon exploration, the latter which led to the *Deepwater Horizon* oil spill in 2010. The effects of these anthropogenic threats on the sperm whale population are poorly understood and little is known about population recovery. In this study, estimated density trends of sperm whales were analyzed from passive acoustic data at three sites in the Gulf of Mexico during and following the *Deepwater Horizon* oil spill over seven years (2010 – 2017). Population structure at each site was taken into account to accurately estimate density trends. Long-term declines were observed at the northern sites, near the *Deepwater Horizon* oil spill and with high shipping presence, suggesting that the decrease could be related to these sources of impact. In contrast, an increase in density was observed at the eastern site, influenced by seasonality. This work represents substantial progress in our understanding of the Gulf of Mexico's sperm whale population, and the potential long-term impact of the *Deepwater Horizon* oil spill as well as other anthropogenic activities. Continuing monitoring of this population can further contribute to more effective conservation strategies.

4.2. Introduction

Assessing species population trends is essential for implementing appropriate management and conservation strategies. Estimating the size of a population is the first step to tracking population trends, yet such estimation remains financially, logistically and conceptually challenging in practice. An increasingly common approach for animal population studies is to employ passive acoustic monitoring (PAM). PAM systems are excellent candidates for long-term and large-scale monitoring and are especially useful to observe remote areas under a range of environmental conditions (e.g., day, night, at depth, under ice, during storms, etc.). Using PAM to estimate the density of marine mammals has gained increasing interest, with recent methods developed to improve the density estimation from single acoustic sensors (von Benda-Beckmann et al., 2018; Frasier et al., 2016; Harris et al., 2018; Küsel et al., 2011; Marques et al., 2013). The most broadly used method for density estimation is distance sampling (Buckland et al., 2001). In distance sampling, modeling the probability of detection from the sensor for each species of interest allows the assessment of animal densities using data from only a subset of the larger area of interest. This requires the quantification of acoustic signals (cues), misclassifications and the corresponding population cue production rates (Buckland, 2006) to convert estimated cue densities to population densities.

Sperm whales have a cosmopolitan distribution and are considered endangered under the Endangered Species Act (ESA) mainly due to commercial whaling which ended in 1988, with the global population size estimated in 2002 at 360,000 animals with estimated densities of 1.4 animals/1000 km^2, approximately one-third of its pre-whaling population size (Whitehead, 2002). Sperm whales are also protected under the Marine Mammal Protection Act (MMPA) and the International Convention for the Regulation of Whaling (IWC). Currently, their population change can be more directly linked to anthropogenic impacts due to an intensifying utilization of the marine environment. Anthropogenic noise and the cumulative risk from multiple stressors including ship strikes, fisheries interactions, oil spills, and pollution are affecting the survival of populations (Farmer et al., 2018). The northern Gulf of Mexico (GOM) is one of the core areas for sperm whales, and was the site of the *Deepwater Horizon* (DWH) oil spill explosion in April 2010, where over 780,000 cubic meters of crude oil were released into the ocean for a total of 87 days (National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling, 2011). The DWH is considered to be the largest offshore oil spill in the petroleum industry's history (Levy and Gopalakrishnan, 2010; Ramseur, 2010). It was followed by a massive cleanup and restoration effort, along with multiple research programs aimed at understanding the effects of this catastrophe, which likely will take decades to achieve.

Sperm whales in the GOM region are considered to be a population or separate "stock" under the MMPA, with population size estimates in 2004 of 1,665 animals and estimated densities of 2.4

animals/1000 km^2 (Whitehead, 2002). In 2009, the minimum population size was estimated to be 763 animals (Hayes et al., 2017). On average, GOM sperm whales are smaller in size and the group size of females and immature animals is about one-third the size of populations found in other areas (Jaquet and Gendron, 2009). The GOM population was historically hunted, with the waters of the Mississippi River reported to be one of the most profitable whaling grounds, and their population characteristics may still be influenced by this depredation (Reeves et al., 2011; Townsend, 1935). The GOM sperm whale population has shown high site fidelity of the core areas identified in the northern GOM, including one area near the Mississippi Canyon where the oil spill occurred (Jochens et al., 2008). A short-term study using PAM near the DWH wellhead before and after the oil spill indicated that some sperm whales moved away from the spill area (Ackleh et al., 2012). However, it is not known if the sperm whale permanently relocated as a result of the spill (Ackleh et al., 2012; Merkens, 2013). Some sperm whales have also been observed in the presence of the oil slicks or contaminated areas (Dias et al., 2017), and likely will continue to remain in the area despite the additional stress imposed by the oil spill. In addition to oil contamination exposure, GOM sperm whales have been exposed to oil spill response activities, such as increased vessel and air traffic, and seismic surveys to assess the wellhead (Dias et al., 2017). The effects of the oil spill on the GOM sperm whale population are poorly understood and little is known about resulting population recovery.

Persistently, the GOM population is exposed to high levels of anthropogenic noise, where their habitat overlaps strongly with heavy shipping traffic lanes, geophysical surveys for hydrocarbon deposits, and large-scale commercial fishery activities (Wright et al., 2007). Furthermore, more hydrocarbon seismic surveys are planned in the following decade with over 4 million km of survey lines in the GOM region (BOEM, Bureau of Ocean Energy Management, 2017). A recent study by Farmer et al. (2018) has linked, through simulation models, a significant reduction in relative fitness of reproductive sperm whale females (4% of the stock reaching terminal starvation) to behavioral disturbance associated with hydrocarbon explorations and substantial decline of up to 25 % of the stock's population.

The long-term effects of the DWH disaster and the anthropogenic noise impact on the population of sperm whales in the GOM are poorly understood due to insufficient precision to estimate population trends. In this study, long-term passive acoustic data collected using near-seafloor High-frequency Acoustic Recording Packages (HARPs) at three sites (**Figure 4.1**) along the continental slope of the GOM region were used to estimate densities of sperm whales between 2010 and 2017. The longest time series to date documenting the estimated densities of sperm whales is provided on a weekly timescale in oiled and unoiled areas during the DWH oil spill, and population trends during the monitored period are evaluated.

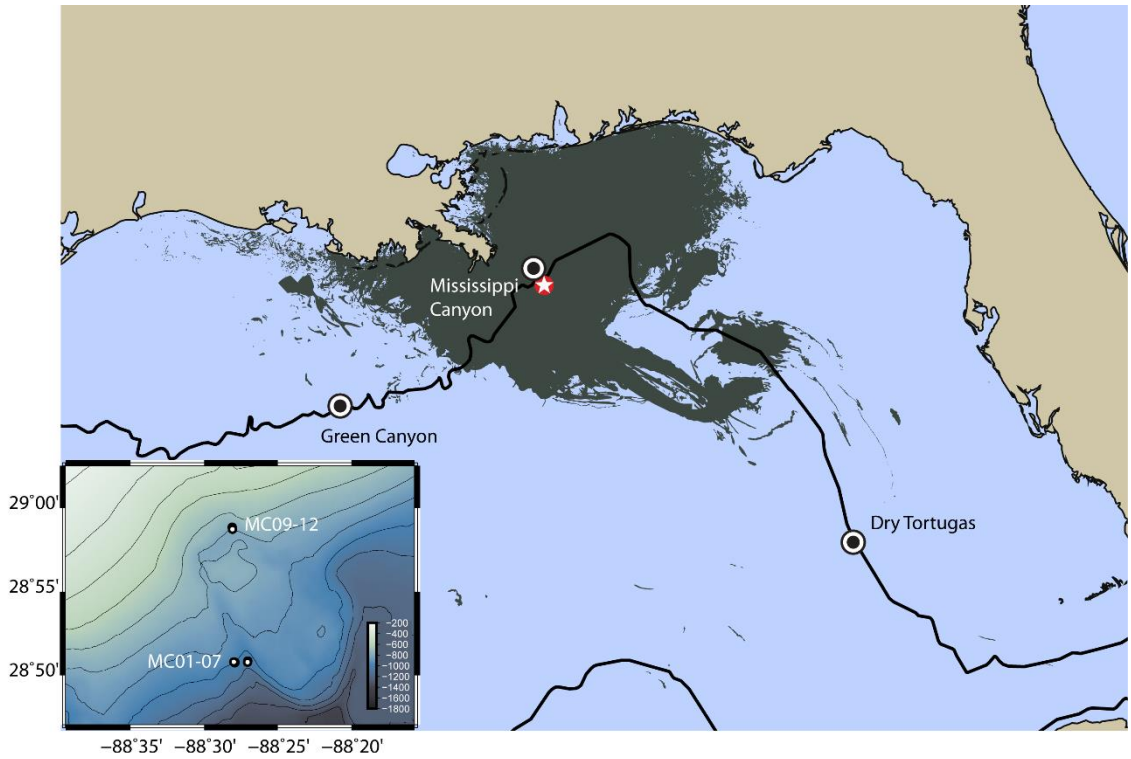


Figure 4.1. Map of deployment locations in the Gulf of Mexico with detections of sperm whales: Green Canyon (GC), Mississippi Canyon (MC), and Dry Tortugas (DT). *Deepwater Horizon* site (red star) and cumulative surface oil during April-August 2010 is shown in dark gray. Surface oil is cumulative NESDIS SAR composite from: http://gomex.erma.noaa.gov. The black line denotes the 1000m contour. Inset shows two alternate deployment locations at the MC site. Map generated using GMT (http://gmt.soest.hawaii.edu/projects/gmt).

4.3. Method

4.3.1. Density estimation

Marques et al. (2009) introduced a method to estimate animal density using acoustic cues detected by fixed passive acoustic sensors. This method has been implemented in numerous studies for estimating densities of different species of marine mammals, using echolocation clicks (Frasier, 2015; Hildebrand et al., 2015; Küsel et al., 2011; Marques et al., 2009; Ward et al., 2012) or calls (Harris et al., 2018; Marques et al., 2011) as cues. By estimating the abundance of detected cues within the monitored area using distance sampling-based methods, the density of cues can be obtained and converted to animal densities using the average cue production rate of the species of interest. The estimated density $\widehat{D}_{kt}$ at site k, during a period of time t is given by

$$\widehat{D}_{kt} = \frac{n_{kt}\,(1-\hat{c}_k)}{\pi\, w^2\;\hat{P}_k\, T_{kt}\,\hat{r}} \qquad (1)$$

where n_{kt} is the number of detected cues at site k and week t, c_k is the proportion of false detections at site k, $\hat{P}_k$ is the estimated averaged probability of detecting a cue within the radius w of the

sensor from the site *k*, T_{kt} represents the total time monitored at site *k* during week *t*, and $\hat{r}$ is the estimated cue production rate.

In this study, regular echolocation clicks of sperm whales from single acoustic sensors of High-frequency Acoustic Recording Packages (HARP) were detected from three deepwater locations in the GOM during and following the DWH oil spill (2010-2017) (**Figure 4.1**). The monitoring locations included a site in Mississippi Canyon near the DWH wellhead (MC), a western site at Green Canyon outside of the DWH surface oil footprint (GC), and an eastern site outside of the oil footprint near the Dry Tortugas (DT). At each site, a deployed HARP recorded sound nearly continuously at a sampling rate of 200 kHz. The acoustic recorder at site MC was moved in location and changed depth. During 2010 – 2013, the hydrophone was deployed at a mean depth of 980 m southwest of a seamount. However, starting in 2014, the hydrophone was deployed 15 km north of the previous location, located north of the seamount at a depth of 800 m.

Automatically detected echolocation clicks were used as cues for estimating weekly densities during weekly intervals at each site. The variance of estimated densities was obtained using the delta method approximation (Seber, 1982):

$$\hat{v}ar\left(\hat{D}_{kt}\right) = \hat{D}_{kt}^2\left\{CV^2(\hat{c}_k) + CV^2\left(\hat{P}_k\right)\right\} \quad (2)$$

where *CV(x)* denotes the coefficient of variation of the random quantity x (i.e., the standard error of the estimate of x divided by the estimate).

4.3.2. Signal description, detection, and classification

Individual sperm whale echolocation clicks were automatically detected using a multi-step approach implemented in MATLAB (Mathworks, Natick, MA). A full description of the algorithms is detailed in **Section 2.3.1.1**. To model the detectability of clicks at different distances, a distribution of source levels must be assumed. In this study, only clicks that exceeded a received level of at least 130 dB_{pp} re 1 µPa were retained for further analyses. This threshold choice was based on a review of histograms of received levels at each site and determined as the received level above which echolocation clicks were reliably detected. Sperm whale encounters were defined as instances with more than 75 s of clicking detected from the same species, with no more than a 30 min gap between successive clicks. The acoustic encounters were then manually reviewed using the custom graphical user interface *DetEdit*, described in **Section 2.3.2**. Misidentified encounters, those from ships or other marine mammals such as beaked whales or delphinids, were removed from the analysis in order to reduce false positives.

4.3.2.1. Detector characterization

Any detection and classification process, including both automated and manual processes is prone to errors. To obtain a reliable estimate, and reduce the bias of density estimates, the detection process must be characterized and accounted for in the estimation (Marques et al., 2013). The effect of noise on the detection process can be attributed to an increase in false alarm rate or a decrease in the detection probability. Therefore, a high amplitude threshold was chosen and periods of time with increased noise, such as shipping passages were considered as no effort time periods due to acoustic masking, where discrimination of sperm whale clicks was not possible. The entire dataset was examined in order to characterize the proportion of false positives at each site using the interactive panels in *DetEdit*, where a subsample of clicks (11 thousand clicks) was randomly selected from the entire dataset and used to assess the presence of false detections.

4.3.3. Probability of detection

The probability of detecting a cue is related to the distance of the vocalizing animal from the sensor, and other factors such as animal diving behavior and orientation to the sensor. Directly calculating distances to the vocalizing animal from single sensors is limited because little spatial information is available. In this case, one must rely on a simulation-based framework (Marques et al., 2013) to estimate the detection probability. The sonar equation can be employed to relate the click received levels to different distances based on signal characteristics, propagation loss, receiver characteristics and sensitivity of the detection system. Monte Carlo simulation methods (Metropolis and Ulam, 1949) have been applied to estimate the probability of detection of numerous vocalizations of marine mammals (Frasier et al., 2016; Harris, 2012; Helble et al., 2013b; Hildebrand et al., 2015; Küsel et al., 2011). By varying the signal and behavioral parameters over many iterations, drawn from information available in the literature, one can incorporate variability and uncertainty into detection probability estimates. Different ways to characterize the detector performance have been implemented, varying from embedding simulated calls in noise, using known source levels at known distances, to modeling distance of clicks based on signal-to-noise ratio, or sound pressure amplitudes (Frasier et al., 2016; Helble et al., 2013a; Hildebrand et al., 2015; Küsel et al., 2011; Ward et al., 2011). The prediction of received levels using this approach requires assumptions about the click characteristics, such as frequency content, beam pattern, and source level that vary as a function of the animal orientation, as well as the animal diving behavior, such as depth distribution or vertical orientation in the water column. Most of these model parameters were derived from acoustic tags and array recordings of sperm whales in the GOM area, and other populations (Irvine et al., 2017; Møhl et al., 2003; Watwood et al., 2006; Zimmer et al., 2005b). Distributions of these parameters were simulated using a Monte Carlo algorithm implemented by Frasier et al., (2016), modified for simulating the detectability of sperm whale echolocation clicks.

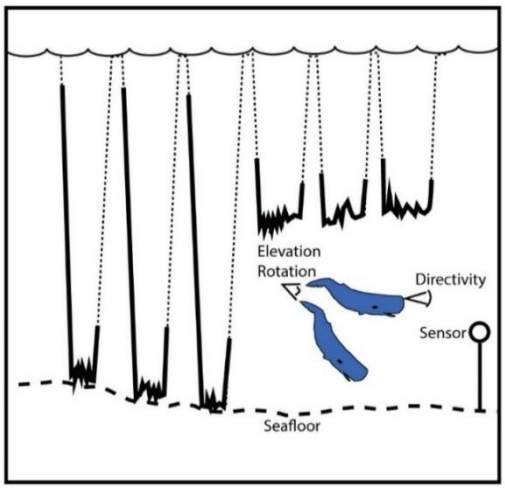


Figure 4.2. Sperm whale click detection model with two modes of diving (dotted lines), near the seafloor (left) and mid-water column (right). The bold portion of the dive track denotes the time spent clicking (following Watwood et al. 2006).

The animal position and orientation with respect to the sensor were simulated (**Figure 4.2**, **Table 4.1**) based on the documented diving behavior of tagged animals in the GOM area (Watwood et al., 2006). The position was modeled in the horizontal plane within a circle of maximum radius of 12 km. A mean dive altitude above seafloor of 300 – 400 m was assumed based on tagged data (Watwood et al., 2006). However, based on diving behavior of tagged animals in the Gulf of California (Irvine et al., 2017), where animals appeared to dive regularly and travel along the seafloor, mean dive altitude was drawn from an asymmetrical bimodal distribution accounting for the proportion of each type of dive reported for the tagged whales, with 85 % of mean dive altitude above seafloor at 300 – 400 m and 15 % at seafloor dive depth (10 m above the seafloor, where the acoustic receiver was deployed). It is assumed that during the acoustic foraging search phase, an animal has certain distributions of depth and orientation. Depth was assigned a uniform distribution between the start depth of clicking (200 m, Watwood et al., 2006) and the target dive altitude, and body orientations were assumed in two dimensions, the vertical (pitch) and horizontal (yaw) planes. A mean body pitch angle of 0º (body parallel to the seafloor) was assumed, with a left-truncated normal, and standard deviation of 0 – 50º during the foraging portions of the dive (Watwood et al., 2006). Because the yaw angle was not documented, and azimuthal symmetry was assumed with respect to the sensor site, all orientations were assumed equally likely (0º – 360º).

The source level and the beam pattern of a click are critical parameters for the orientation-dependent transmission loss. The transmission loss indicates if a click would be detectable within the maximum range, and this was subtracted from the on-axis source level as

$$RL = SL - TL \tag{3}$$

where the sonar equation relates the click received levels (RL) with the modeled transmission loss (TL) and the effective on-axis source level (SL). Source level of on-axis clicks of male sperm whales from the Mediterranean and Norway have been reported to be at least 245 ± 3 dB_{pp} re: 1µPa at 1m with a directivity of 27 dB (Møhl et al., 2003; Zimmer et al., 2005b). Since the sperm whale GOM population is composed mostly of adult females and immature individuals the source level for this population was expected to be lower, around 230 ± 3 dB dB_{pp} re 1µPa at 1m and directivity of 27 dB, based on the animals' smaller size and within the expected correlation of source levels and body mass (Hildebrand et al., in review; Jensen et al., 2018). Signal peak frequency from adult males has been reported to be 12 kHz (Møhl et al., 2003; Zimmer et al., 2005b), and knowing that it is affected by sound absorption, the peak frequency was approximated by taking the mean of the peak frequencies of the detected clicks. The measured mean click peak frequency was 11 kHz and was used to approximate transmission loss. The resulting attenuation of source levels as a function of off-axis angle with respect to the sensor was estimated from beam axes reported by the literature. The model of a circular piston (Zimmer et al., 2005b) has been used to estimate the acoustic beam pattern of sperm whales (Møhl et al., 2003; Zimmer et al., 2005b) and other cetacean species (Frasier et al., 2016; Hildebrand et al., 2015; Küsel et al., 2011; Rasmussen et al., 2004; Zimmer et al., 2005a). The off-axis beam shape was completed by interpolating the two distinct beam patterns of sperm whale clicks reported by Zimmer et al. (2005b), where click levels extend both in the forward and backward directions. The off-axis angles were drawn from a uniform distribution with the minimum (35 – 40 dB) and maximum (25 – 35 dB) amplitudes reported in the literature (Zimmer et al., 2005b) (**Table 4.1**).

Transmission loss was simulated, as in Frasier et al. (2016), using the ray-tracing algorithm Bellhop (Porter and Bucker, 1987), with site-specific environmental and physical parameters drawn from the Oceanographic and Atmospheric Master Library (OAML). The model simulation involved placement of 10,000 animal positions vocalizing within a 12 km radius of the sensor, with varying parameter values assigned to each source. Clicks were considered to be detected if their received levels were at or above 130 dB_{pp} re: 1µPa, mimicking the detection threshold used in the signal analysis. The detection probability associated with each simulation (500 model iterations) was computed as the ratio of clicks detected to the total number of simulated clicks, and a mean probability of detection and its variance was derived from the Monte Carlo framework (**Table 4.1**).

4.3.4. Cue production rate

To calculate animal density from passive acoustics, one must know the species-specific click production rate, which is the average number of cues produced per animal per time. Cue rates are used to translate the number of cues produced to the number of animals producing them. The cue rate can vary significantly between populations, especially for sperm whales where population

Table 4.1. Literature-based signal and behavior parameters used in Monte Carlo simulation of diving sperm whale detectability.

Parameter	Units	Mean (μ)	Standard Deviation (σ)	Distribution	References
Dive depth	Meters %	300 – 400 85 – 95	10 – 30 1 – 2	Asymmetrical bimodal	1, 2
Benthic dive depth	Meters %	Sea floor 5 – 15	10 – 20 1 – 2		
Start clicking depth	Meters	200 – 220	65 – 75	None	1
Orientation: Elevation	Degrees	0	0 – 50	Normal, left truncated at 0º	1
Orientation: Azimuth	Degrees	0 – 359	n/a	Uniform	Simulation assumption
Source level	dB_{pp}	225 – 235	2 – 5	Normal	Simulation assumption
Directivity	dB	25 – 30	n/a	Uniform	3, 4
90º off-axis TL	dB	35 – 40	n/a	Uniform	4
180º off-axis TL	dB	25 – 30	n/a	Uniform	4
Peak frequency	kHz	11	n/a	None	This dataset
Maximum detection range	Km	20	n/a	None	This dataset

[1.] Watwood et al. 2006
[2.] Irvine et al. 2017
[3.] Møhl et al. 2003
[4.] Zimmer et al., 2005b

structure varies in relation to the latitude, with females and immature present in low latitudes, and males between tropic and higher latitudes (Connor et al., 1998). Sperm whales are significantly sexually dimorphic, thus influencing the acoustic behavior of the sexes. Adult males and females are known to have different clicking rates, with males clicking nearly every 0.85 – 1 s, and females every 0.5 s (Goold and Jones, 1995). It is critical to ensure that the cue rate estimate is representative of the population under study. Therefore, in the case of sperm whales, it is crucial to account for the population structure. Ideally, cue rate would be estimated from the animals being studied, at the same time and place as the acoustic survey. However, when using single acoustic sensors, animals are only known to be present when their clicks are detected. To convert the number of animals vocalizing into a total number of animals present, an estimate of the proportion of time of an animal spent vocalizing was required. Consequently, animals diving behavior must be obtained from auxiliary data. Cue rates were obtained here from the product of the mean proportion of time an animal was clicking, and the inverse of the inter-click interval (ICI) of a series of clicks. The ICI was estimated from the acoustic recordings, while the diving characteristics were obtained from tag data reported mostly from animals in the GOM area (Irvine et al., 2017; Watwood et al., 2006). The diving and foraging behavior has been studied for geographically diverse populations of sperm whales (Ward et al., 2012; Watwood et al., 2006),

and has been found to be remarkably similar. A dive cycle is defined as the time when an animal is clicking at a regular rate during a deep dive plus the silent times spent on the surface. For the studied populations, sperm whales on average spend greater than 72% of their time in foraging dive cycles, producing regular clicks for approximately 81% of a dive (Watwood et al., 2006). Tagged sperm whales in the GOM, were primarily females and immature whales, with dive cycles reported for this sex class of 45.5 min deep-dives, and 8.1 min surface intervals, with animals spending 81.2 % of the dive time in search phase producing regular clicks. The proportion of the dive cycle spent clicking was calculated as the proportion of time spent in search phase during a deep dive and the total time of a dive cycle.

To account for the population structure of sperm whales in the area under study, the detected encounters were manually categorized into likely sexes based on ICI distributions. The development and implementation of this method is described in **Section 3.3**. Distinct and consistent ICI distributions were found in the dataset allowing the differentiation of adult males from mixed groups of whales, consisting of adult females and their offspring. When animals were close to the sensor, multiple animals were detected, or when distant from the sensor, occasionally intermediate clicks were not detected, thus showing a less characteristic clicking rate. Therefore, it was determined that the mode of the ICI distribution was the most robust way of estimating the characteristic cue rate of the population. The modal ICI for both adult males and the mixed group were obtained, and weighted by the proportion of presence of both classes, resulting in a weighted modal ICI.

4.3.5. Densities trend analysis

Long-term trends of sperm whale weekly densities were estimated for each site using a Thiel-Sen regression with 5 – 95 % confidence intervals obtained using a bootstrap method. The median slope across 500 pairs of points selected randomly with replacement within each time series was computed 100 times. As described in **Section 3.4**, seasonal presence of sperm whales was found in the monitoring sites. The seasonal component was removed from the regression using a monthly seasonal trend decomposition procedure (Cleveland et al., 1990). Theil-Sen slope estimates for linear trends in the time series were computed for the deseasoned data as in Frasier, (2015), and percentage of change per year at the three sites were computed.

4.4. Results

Seven years (2010 – 2017) of passive acoustic monitoring data were available during and following the DWH oil spill in the GOM. In total, 6,438 h of data (202 TB) were analyzed, over which a total of 34 million sperm whale clicks were detected. Sperm whales were detected at all sites with 71 % of the total detected clicks at site MC, 24 % at site GC, and 5 % at site DT. Mean false positive rate at site GC was the lowest (3 %) of all sites, followed by site MC when acoustic

Table 4.2. Average sperm whale densities per site, GC, MC (including both depths), DT (including seasonality) given in # of animals per 1000 km^2. Parameters used for density estimation include the average number of clicks per second n_{kt}/T_{kt}, the percentage of false clicks c_k with associated CV, the expected cue rate r with associated CV, the maximum horizontal range w, and the probability of detection P_k with associated CV.

Site	Density (#/1000km^2) ± stdev		N_{kt}/T_{kt} (#/sec)	c_k (% False clicks)	CV	r Cue Rate (#/sec)	CV	w Max range (km)	P_k Prob. detect	CV
GC	1.28	± 0.81	0.05	2.7	0.23	1.45	0.35	12	0.05	0.49
MC_{980}	1.84	± 1.12	0.09	6.7	0.14	1.51	0.35	12	0.07	0.48
MC_{800}	1.26	± 0.77	0.06	3.4	0.20	1.45	0.35	12	0.06	0.46
DT	0.34	± 0.27	0.01	8.2	0.52	1.35	0.35	12	0.05	0.50
$DT_{in\ seas.}$	0.59	± 0.47	0.02	8.2	0.52	1.37	0.35	12	0.05	0.49
$DT_{out\ seas.}$	0.15	± 0.12	0.01	8.2	0.52	1.28	0.35	12	0.05	0.49

recorder was deployed at 800 m depth (**Table 4.2**). Although at sites DT and MC (during the deployment at depth 980 m) the mean false positive rate was higher than at GC, i.e. 8 % and 7 %, respectively, overall, mean false positive rates were low.

4.4.1. Detection Probability

Received levels from detected clicks were compared to the predicted received levels of the simulation model (**Figure 4.3**). Received levels were similar between the detected and predicted clicks, with fewer detected clicks of high amplitude with respect to those predicted by the model. Site DT had more error in the fit between the detected and the predicted received levels than either GC or MC. Poor fit at high received levels for all sites is most likely due to low numbers of high amplitude clicks due to finite recording intervals. As strong seasonality was observed at site DT (see **Section 3.4.3**), detection probabilities were also calculated during season and off-season, and compared to the probability of detection when not accounting for seasonality (**Table 4.2**).

The detection probabilities were similar for all three sites (**Table 4.2**), suggesting little or no dependence upon local site features. Detectability dropped off rapidly for all sites at ranges between 1 and 2.5 km from the sensor (**Figure 4.4**), owing to the highly directional beam pattern of sperm whale clicks. Because the maximum range of detectability was estimated at a range of 12 km, detection probabilities were estimated for both MC locations. However, no significant difference was found between the two locations.

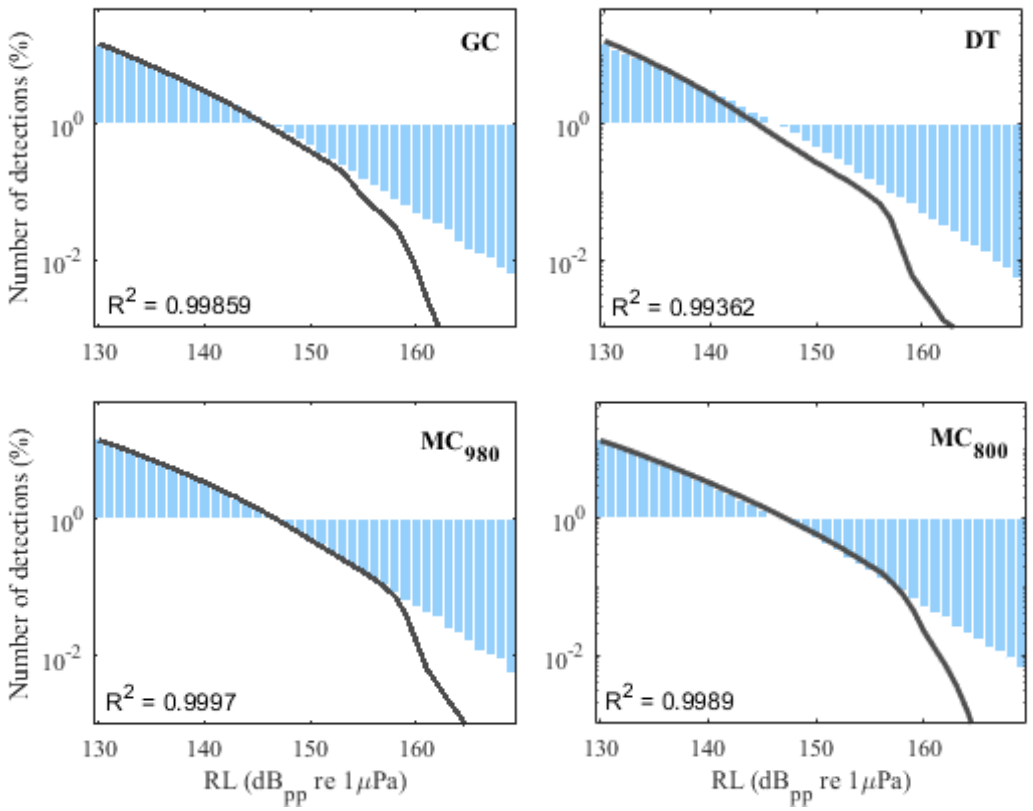


Figure 4.3. Comparison of percentage of received levels (RL) of detected clicks in logarithmic scale (black line) and the model predicted RL (blue bars) for the three sites, GC, DT, and MC (including both deployments at 980 and 800 m depth).

4.4.2. Cue production rate

All encounters that were classified as mixed group or adult males were used to estimate cue rates per site, by obtaining the product of the mean proportion of the dive cycle spent clicking and the inverse of the weighted modal ICI (**Table 4.3**). ICIs per site and the sex class were used to calculate the weighted modal ICI per site. Site DT had the highest proportion of adult males, yielding a slightly lower weighted modal ICI than the other sites. Estimated cue rates were 1.5 clicks/s, with a slightly lower rate at site DT of 1.4 clicks/s. During the off-season at DT, the proportion of adult males increased notably, yielding a lower estimated cue rate (1.3 clicks/s).

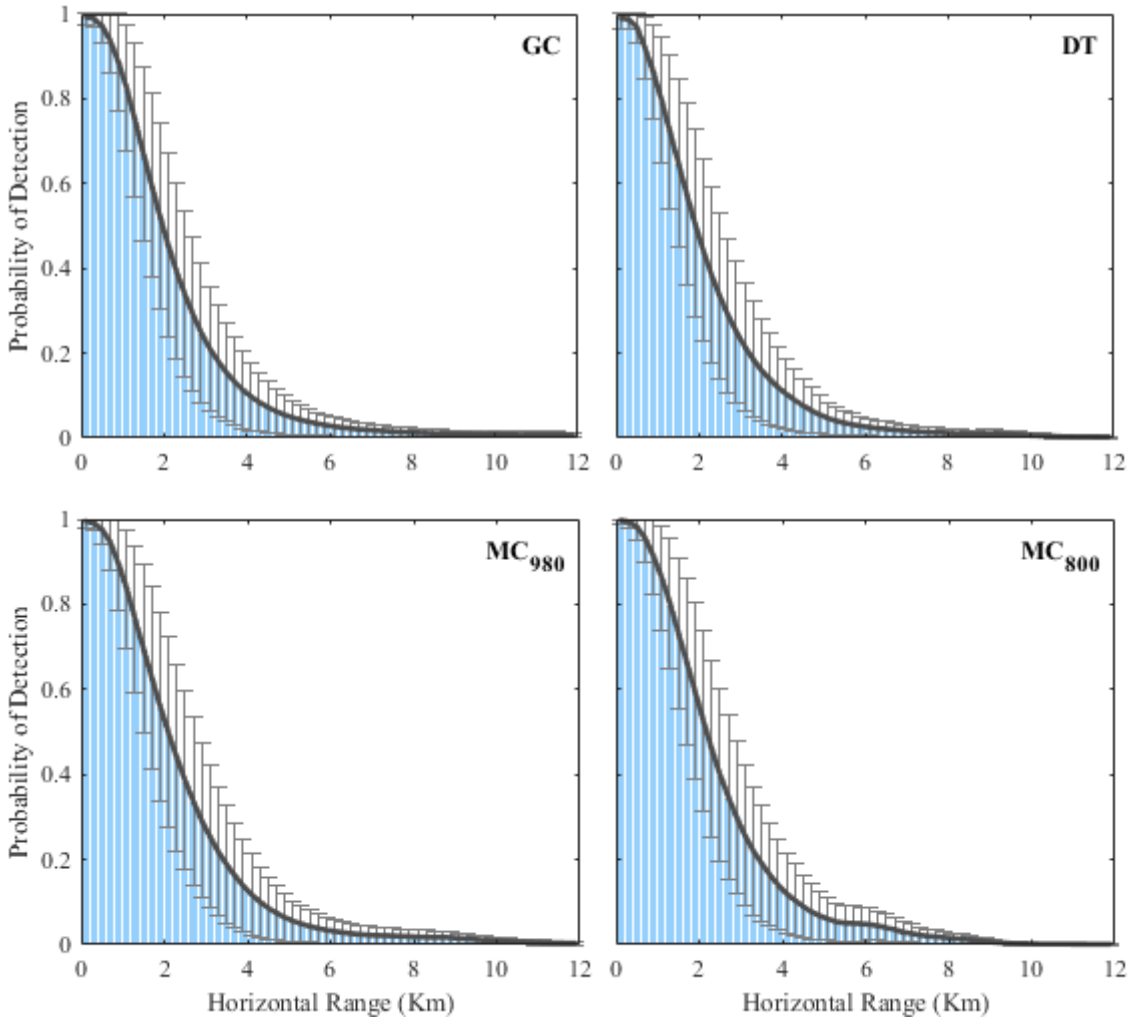


Figure 4.4. Estimated detection probability for sperm whale clicks based on a simulation using sound propagation modeling for site GC, DT, and MC (including both mean depth at 980 and 800 m). Error bars represent 1 standard deviation from the mean.

Table 4.3. Weighted modal inter-click interval (ICI) for sperm whale at GOM recording sites GC, DT, and MC (including both mean depth). Modal ICI per each sex class and ratio of each class used to calculate the weighted modal ICI.

Site	Modal ICI (ms)		Ratio (%)		Weighted Modal ICI (ms)
	Mixed group	Adult males	Mixed group	Adult males	
GC	473.5	756.6	98.8	1.2	476.9
MC_{980}	455.0	811.7	99.8	0.2	455.6
MC_{800}	475.1	797.8	99.9	0.1	475.6
DT	500.1	811.0	96.8	3.2	510.0
$DT_{in\ seas.}$	499.4	805.3	98.4	1.6	504.2
$DT_{out\ seas.}$	508.5	809.7	90.4	9.6	537.5

4.4.3. Density estimates and trends

Higher densities were observed at the northern sites (MC and GC) than the eastern site (DT) (**Figure 4.5, Table 4.2**), with estimated densities for site GC, MC, and DT of 1.28 ± 0.85, 1.26 ± 0.77, and 0.33 ±0.27 animals/1000 km^2, respectively. At site MC, sperm whales were present consistently during most of the monitored weeks. At site GC, whale presence fluctuated among

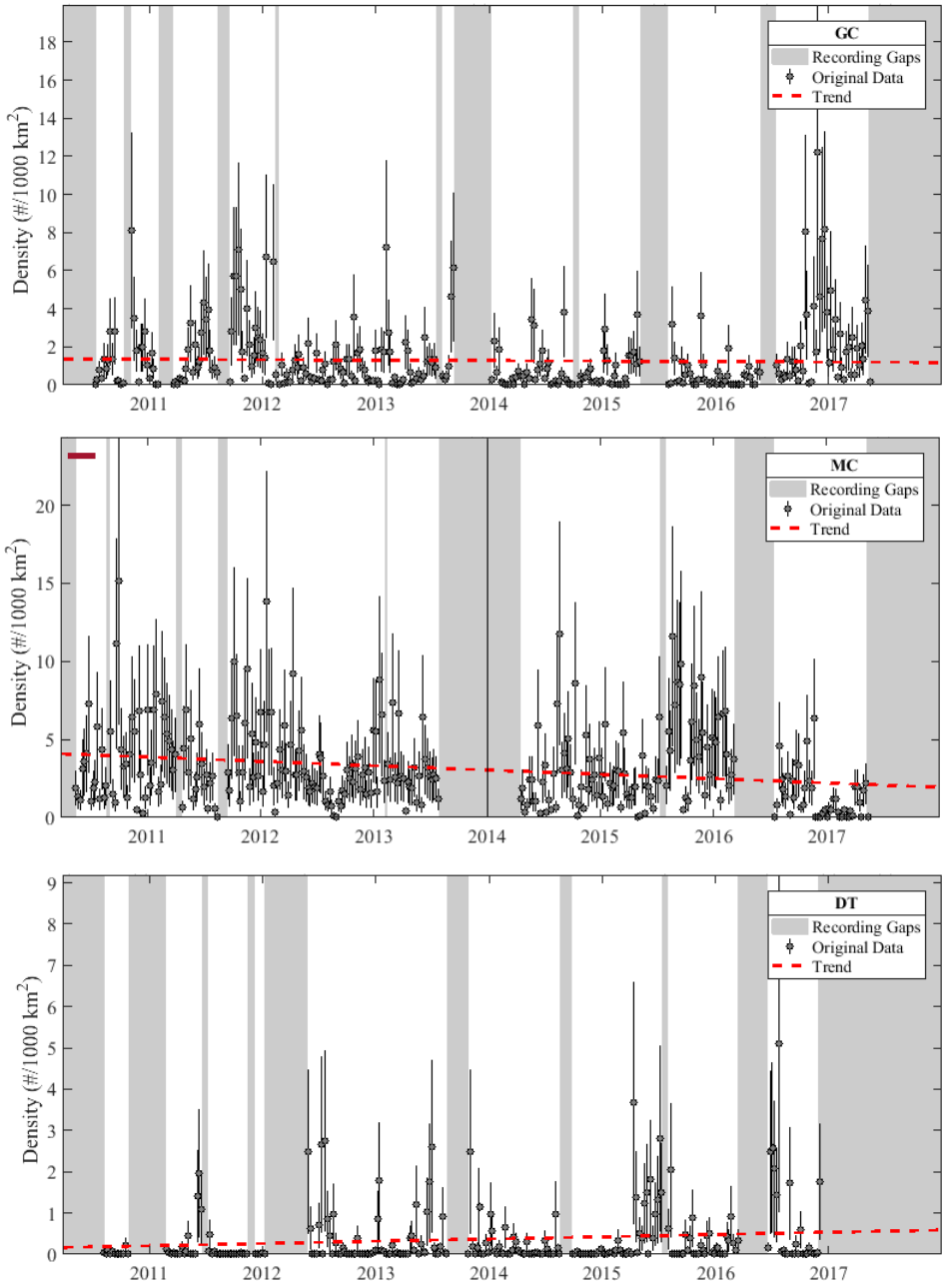


Figure 4.5. Weekly density estimates of sperm whales at site GC, MC, and DT. Circles denote estimates and vertical lines show +/- one standard error. Red dashed line shows the de-seasonal Theil-Sen trend of densities along monitoring time. Shaded areas lack recording effort. Middle plot: black line indicates change of deployment location from 980 m to 800 m depth; red line indicates time of the DWH oil spill.

the monitored years. Whale presence varied seasonally at site DT with animals present primarily in the summer months (April-August). Densities at site DT derived from parameters computed from the entire time series varied slightly from densities estimated accounting for the seasonality differences, with 0.3 ± 0.3 animals/1000 km^2 to 0.2 ± 0.1 animals/1000 km^2, respectively.

Long-term trend estimates of weekly densities suggest a slow decline at the western sites between 2010 and 2017, with a greater decline at site MC with -6.8 % (95% CI [-8.2, -5.2]) annual change, and -2.5 % (95% CI [-5.7, 1.1]) at site GC. In contrast, an increase in density was found at the eastern site with an annual change of 32.5 % (95% CI [13.8, 79.1]). However, overall densities were the lowest at this site.

4.4.4. Shipping presence

Detected shipping passages were excluded from the analysis (**Figure 4.6**). Shipping presence was higher at the northern sites (GC and MC), with the noisiest times following the oil spill (April-August, 2010) at site MC, with ship presence above 20 h per week. However, the most consistent and noisiest periods were at site GC between March 2012 and May 2016, where presence doubled from the earliest years. The eastern site (DT) had the least shipping, with only two short periods with shipping presence reaching 20 h per week.

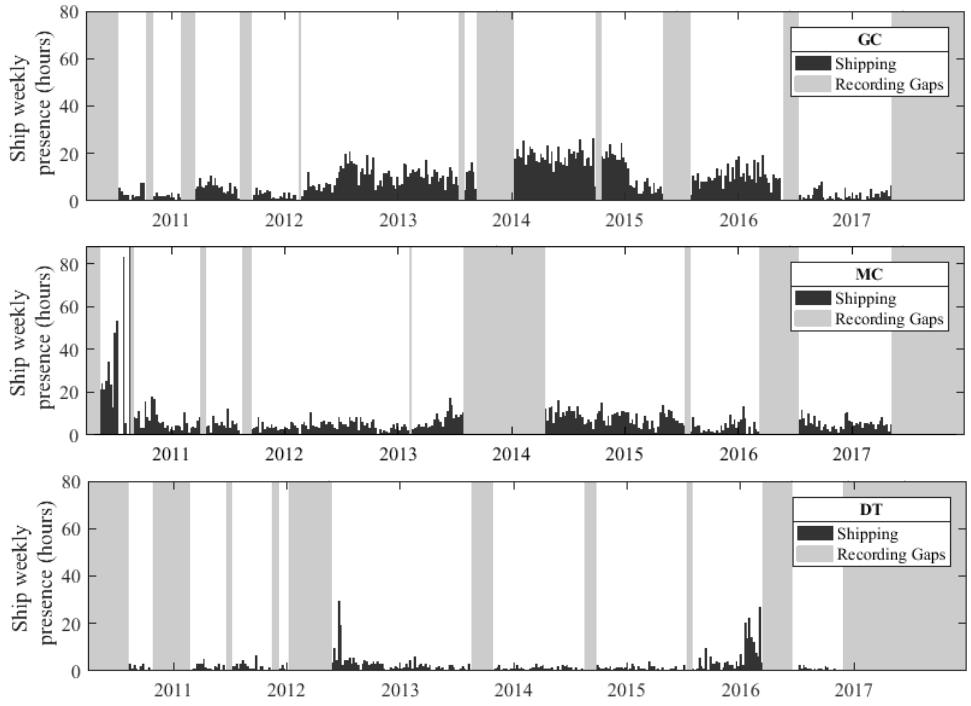


Figure 4.6. Weekly presence of shipping as number of hourly bin with detections at site GC, MC, and DT. Shaded areas lack recording effort.

4.5. Discussion

The key advantage of PAM in comparison to visual and aerial studies is the ability to collect continuous and long-term data from populations, especially for species with low probability of visual detection. PAM may be the most effective way to assess the population trends of sperm whales, which on average spend greater than 72 % of their time submerged (Watwood et al., 2006). However, PAM is limited in its ability to determine absolute population abundance due to the restricted spatial coverage provided by fixed acoustic recorders (Hildebrand et al., 2015). In this study, estimated densities specific to the monitoring sites were provided from single acoustic sensors, applying a cue-based distance sampling method (Marques et al., 2013). This approach required modeling the detection probability of cues within a maximum radius of 12 km around the sensor. Detection probabilities were high near the sensor, and declined rapidly between 1 and 2.5 km, resulting in low overall detectability beyond 8 km radius. The overall detectability was set by the high amplitude threshold chosen for the signal analysis (130 dB_{pp} re 1 µPa), which was necessary to ensure that clicks were reliably detected within the defined area. Most detected clicks were received off-axis because the high signal directivity makes the probability of detecting an on-axis click very low. Similarities between predicted and detected RL distributions suggest that this was the case for our recordings. Although substantial similarity was observed for most sites, the eastern site (DT) had the most error between detected and predicted RL. Perhaps the greater depth of the sensor at this site requires adjustment of the model to account for diving behavior in its vicinity. Detectability was also compared between on and off season, yielding similar results. One assumption made for the simulation was that animals were uniformly distributed around the sensor. In the future, more complex spatial distributions, such as a non-uniform distribution, could be added to the simulation framework with better information about their distribution within the area. Another potential source of error is the simulation of the beam pattern, where depth-dependent variation at ranges larger than 6 km (Zimmer et al., 2005b) could have created bias. Additionally, modeling of acoustic propagation could introduce error. Transmission loss was approximated using a single peak frequency value (11 kHz). Ignoring the frequency dependence and seasonal water conditions, such as temperature and salinity, when modeling the transmission loss could have resulted in an underestimation of the effective detection area (von Benda-Beckmann et al., 2018). Furthermore, dive depth was simulated as an asymmetrical bimodal distribution to account for the benthic dives observed for other populations of sperm whales (Irvine et al., 2017). Accurate values of the proportion of these types of dives for the studied population could vary the detectability if, for example, more dive depths were near to the depth of the fixed bottom-mounted sensors.

Although there is detailed information about the diving and vocal behavior, and click characteristics of sperm whales, particularly for the GOM population (Irvine et al., 2017; Mate et

al., 2017; Møhl et al., 2003; Watwood et al., 2006; Zimmer et al., 2005b), no studies have accounted for population structure when estimating densities of marine mammals. Sperm whales have stratified distributions of sexes with matrilineal groups mainly found in the tropics and sub-tropics, and adult males in higher latitudes (Connor et al., 1998), suggesting that parameters needed for density estimation would likely be spatially specific (Douglas et al., 2005). Therefore, based on the observed population structure in the monitored area (see **Section 3.4**), we approximated the cue production rates for each of the monitored sites. Different proportions of mixed groups (matriarchal groups) and adult males were observed among the study sites, with a higher proportion of adult males at the eastern site (DT), especially during the off season. Cue rates were estimated from the weighted modal ICI with the proportion of sexes, and ICI observed from the recordings. This resulted in slightly different cue rates, which allowed for more accurate density estimation at each site. However, we have little to no data on the differences in diving behavior and click characteristics between sexes in this region. Source levels and beam patterns of sperm whale clicks are known from adult males of other populations (Møhl et al., 2003; Zimmer et al., 2005b), and no information exists for smaller animals, such as adult females and immature whales. Lower source levels and directionalities for the corresponding population were assumedwhich is mostly formed of adult females and immature whales (**Figure 4.3**, **Table 4.1**). Source level is critical to estimate the detection range. Therefore, the source level and beam pattern assumptions made in this study could have significantly influenced the detection probabilities. Moreover, diving behavior used for the simulation is based on tagged females and immature whales from the GOM area (Watwood et al., 2006), and likely biased the detectability in cases where there is a significant proportion of adult male presence. Future work may seek to improve click simulation to allow for more refined estimates of click detectability with a more complex behavioral model than the one implemented here.

Comparing the density estimates obtained here and the existing estimates based on visual surveys provided further insight on the accuracy of the obtained numbers. The NOAA estimates provided for global population is of 1.4 animals/1000 km^2, with 2.3 animals/1000 km^2 in the GOM area (Whitehead, 2002). The mean density estimated from this study is of maximum 1.3 animals/1000 km^2, which supports the number reported for the global population, but is lower than the densities that have been previously reported for the GOM area. Higher densities were observed at the northern sites than the eastern site, and different trends of densities were observed among the three monitored sites during the seven year period. Population trends at site MC may be related to exposure to the DWH oil spill, where animals most likely interacted with the surface and subsurface oil for extended periods of time in the months following the oil spill. Sperm whale densities declined overall during the study period in the northern sites (MC and GC), while increasing at the eastern site (DT), where lower densities were observed relative to the northern

sites. Since the strongest declines were seen at site MC near the DWH oil spill, population declines could be related to the DWH event. However, densities at site MC were the highest of the three monitored sites, providing evidence that this area is core habitat (Jochens et al., 2008). Declines at the northern sites also could be related to shipping presence or natural inter-annual variability. Declines at site GC between 2012 and 2016 correlate with a strong increase in shippping presence during the same period of time. This may indicate that population declines at this site may be related to noise impacts or other shipping impacts. To a lesser extent, the increase in noise due to shipping presence at site MC may also have contributed to the decline in whales observed there. However, it is not possible to confidently interpret the density increase at site DT, because of a strong seasonality. A broader understanding of sperm whale response to local conditions due to both natural and anthropogenic processes are needed to intepret the site-level trends. Investing in continued monitoring to collect longer time series of acoustic data will potentially yield to the abilities to identify the drivers of this trend.

4.6. Conclusions

Passive acoustic data were analyzed to estimate the density of sperm whales at three sites in the Gulf of Mexico during and following the DWH oil spill. A Monte Carlo simulation framework was used to estimate the probability of detecting sperm whales using a click-based density estimate method. To reduce potential errors in density estimates related to vocalization rates, population structure at each site was taken into account to more accurately estimate cue production rates. Potential impacts of the DWH oil spill and shipping presence on the sperm whale population were evaluated by calculating long-term trends over the density times series, and it was suggested that the decrease at the northern sites could be related to these sources of impact. Contrastingly, an increase of sperm whale densities was observed at site DT. However, was not possible to relate the increase at this site with the decreases seen at the other sites because of a strong seasonal presence which requires a better understanding of the local conditions to determine the site trend.

4.7. Acknowledgments

Funding for HARP data collection and analysis was provided by the Natural Resource Damage Assessment partners (20105138), the US Marine Mammal Commission (20104755/E4061753), the Southeast Fisheries Science Center under the Cooperative Institute for Marine Ecosystems and Climate (NA10OAR4320156) with support through Interagency Agreement #M11PG00041 between the Bureau of Offshore Energy Management, Environmental Studies Program and the National Marine Fisheries Service, Southeast Fisheries Science Center, and the CIMAGE Consortium of the Gulf of Mexico Research Initiative (SA 12-10/GoMRI-007). I thank LenThomas, Tiago Marques, Danielle Harris, Steve Murawski and Sheryl Gilbert for conveying the knowledge to the statistical methods used for this project.

4.8. References

Ackleh, A. S., Ioup, G. E., Ioup, J. W., Ma, B., Newcomb, J. J., Pal, N., Sidorovskaia, N. A., et al. (**2012**). "Assessing the Deepwater Horizon oil spill impact on marine mammal population through acoustics: Endangered sperm whales," J. Acoust. Soc. Am., **131**, 2306. doi:10.1121/1.3682042

von Benda-Beckmann, A. M., Thomas, L., Tyack, P. L., and Ainslie, M. A. (**2018**). "Modelling the broadband propagation of marine mammal echolocation clicks for click-based population density estimates," J. Acoust. Soc. Am., **143**, 954–967. doi:10.1121/1.5023220

BOEM (Bureau of Ocean Energy Management) (**2017**). *Gulf of Mexico OCS proposed geological and geophysical activities: Western, Central, and Eastern planning areas.* Final Environmental Impact Statement, OCS EIS/EA: BOEM 2017-051.

Buckland, S. T. (**2006**). "Point-Transect Surveys for Songbirds: Robust Methodologies," Auk, **123**, 345. doi:10.1642/0004-8038(2006)123[345:PSFSRM]2.0.CO;2

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (**2001**). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*, Oxford University Press, 432 pages. Retrieved from https://books.google.com/books?id=9-LiOqoKE9QC&pgis=1

Cleveland, R. B., McRae, W. S., and Terpenning, I. (**1990**). "STL: A seasonal-trend decomposition procedure based on loess," J. Off. Stat., **6**, 3–33.

Connor, R. C., Mann, J., Tyack, P. I., and Whitehead, H. (**1998**). "Social evolution in toothed whales," Trends Ecol. Evol., **13**, 228–232.

Dias, L. A., Litz, J., Garrison, L., Martinez, A., Barry, K., and Speakman, T. (**2017**). "Exposure of cetaceans to petroleum products following the Deepwater Horizon oil spill in the Gulf of Mexico," Endanger. Species Res., **33**, 119–125. doi:10.3354/esr00770

Douglas, L. A., Dawson, S. M., and Jaquet, N. (**2005**). "Click rates and silences of sperm whales at Kaikoura, New Zealand," J. Acoust. Soc. Am., , doi: 10.1121/1.1937283. doi:10.1121/1.1937283

Farmer, N. A., Baker, K., Zeddies, D. G., Denes, S. L., Noren, D. P., Garrison, L. P., Machernis, A., et al. (**2018**). "Population consequences of disturbance by offshore oil and gas activity for endangered sperm whales (Physeter macrocephalus)," Biol. Conserv., **227**, 189–204. doi:10.1016/j.biocon.2018.09.006

Frasier, K. E. (**2015**). *Density estimation of delphinids using passive acoustics: A case study in the Gulf of Mexico.* Ph.D. Thesis, University of California San Diego, La Jolla, CA, 262

pages.

Frasier, K. E., Wiggins, S. M., Harris, D., Marques, T. A., Thomas, L., and Hildebrand, J. A. (**2016**). "Delphinid echolocation click detection probability on near-seafloor sensors," J. Acoust. Soc. Am., **140**, 1918–1930. doi:10.1121/1.4962279

Goold, J. C., and Jones, S. E. (**1995**). "Time and frequency domain characteristics of sperm whale clicks," J. Acoust. Soc. Am., **98**, 1279–1291. doi:10.1121/1.413465

Harris, D. V. (**2012**). *Estimating whale abundance using sparse hydrophone arrays.* University of St Andrews, St. Andrews, UK, 10-43 pages.

Harris, D. V., Miksis-Olds, J. L., Vernon, J. A., and Thomas, L. (**2018**). "Fin whale density and distribution estimation using acoustic bearings derived from sparse arrays," J. Acoust. Soc. Am., **143**, 2980–2993. doi:10.1121/1.5031111

Hayes, S. A., Josephson, E., Maze-Foley, K., and Rosel, P. E. (**2017**). *US Atlantic and Gulf of Mexico Marine Mammal Stock Assessments -- 2016* NOAA Tech Memo NMFS NE,166 Water Street, Woods Hole, MA 02543-1026, 274 pages. doi:10.7289/V5/TM-NEFSC-241

Helble, T. A., D'Spain, G. L., Campbell, G. S., and Hildebrand, J. A. (**2013**). "Calibrating passive acoustic monitoring: Correcting humpback whale call detections for site-specific and time-dependent environmental characteristics," J. Acoust. Soc. Am., **134**, EL400-EL406. doi:10.1121/1.4822319

Helble, T. A., D'Spain, G. L., Hildebrand, J. A., Campbell, G. S., Campbell, R. L., and Heaney, K. D. (**2013**). "Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones," J. Acoust. Soc. Am., **134**, 2556–2570. doi:10.1121/1.4816581

Hildebrand, J. A., Baumann-Pickering, S., Frasier, K. E., Trickey, J. S., Merkens, K. P., Wiggins, S. M., McDonald, M. A., et al. (**2015**). "Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico," Sci. Rep., **5**, 16343. doi:10.1038/srep16343

Hildebrand, J. A., Frasier, K. E., Baumann-pickering, S., Wiggins, S. M., Merkens, K. P., Garrison, L. P., Soldevilla, M. S., et al. (n.d.). "Assessing the presence of pygmy and dwarf sperm whales in the Gulf of Mexico using passive acoustic monitoring," Front. Mar. Sci.,.

Irvine, L., Palacios, D. M., Urbán, J., and Mate, B. (**2017**). "Sperm whale dive behavior characteristics derived from intermediate-duration archival tag data," Ecol. Evol., , doi: 10.1002/ece3.3322. doi:10.1002/ece3.3322

Jaquet, N., and Gendron, D. (**2009**). "The social organization of sperm whales in the gulf of California and comparisons with other populations," J. Mar. Biol. Assoc. United Kingdom,

89, 975–983. doi:10.1017/S0025315409001507

Jensen, F. H., Johnson, M., Ladegaard, M., Wisniewska, D. M., and Madsen, P. T. (**2018**). "Narrow Acoustic Field of View Drives Frequency Scaling in Toothed Whale Biosonar," Curr. Biol., , doi: 10.1016/J.CUB.2018.10.037. doi:10.1016/J.CUB.2018.10.037

Jochens, A., Jochens, A., Biggs, D., Biggs, D., Benoit-Bird, K., Benoit-Bird, K., Engelhaupt, D., et al. (**2008**). "Sperm whale seismic study in the Gulf of Mexico Synthesis Report," Mms, **96**, 3268. doi:10.1121/1.410971

Küsel, E. T., Mellinger, D. K., Thomas, L., Marques, T. a, Moretti, D., and Ward, J. (**2011**). "Cetacean population density estimation from single fixed sensors using passive acoustics," J. Acoust. Soc. Am., **129**, 3610–3622. doi:10.1121/1.3583504

Levy, J. K., and Gopalakrishnan, C. (**2010**). "Promoting Ecological Sustainability and Community Resilience in the US Gulf Coast after the 2010 *Deepwater Horizon* Oil Spill," J. Nat. Resour. Policy Res., **2**, 297–315. doi:10.1080/19390459.2010.500462

Marques, T. a., Munger, L., Thomas, L., Wiggins, S., and Hildebrand, J. a. (**2011**). "Estimating north pacific right whale Eubalaena japonica density using passive acoustic cue counting," Endanger. Species Res., **13**, 163–172. doi:10.3354/esr00325

Marques, T. a., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. a., Moretti, D. J., Harris, D., et al. (**2013**). "Estimating animal population density using passive acoustics," Biol. Rev., **88**, 287–309. doi:10.1111/brv.12001

Marques, T. a, Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. L. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales," J. Acoust. Soc. Am., **125**, 1982–94. doi:10.1121/1.3089590

Mate, B. R., Irvine, L. M., and Palacios, D. M. (**2017**). "The development of an intermediate-duration tag to characterize the diving behavior of large whales," Ecol. Evol., **7**, 585–595. doi:10.1002/ece3.2649

Merkens, K. (**2013**). *Deep-Diving Cetaceans of the Gulf of Mexico: Acoustic Ecology and Response to Natural and Anthropogenic Forces Including the Deepwater Horizon Oil Spill* University of California, San Diego, 146 pages.

Metropolis, N., and Ulam, S. (**1949**). "The Monte Carlo Method," J. Am. Stat. Assoc., **44**, 335–341. doi:10.1080/01621459.1949.10483310

Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., and Lund, A. (**2003**). "The monopulsed nature of sperm whale clicks," J. Acoust. Soc. Am., **114**, 1143–1154. doi:10.1121/1.1586258

National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling (**2011**). *National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling*.

Porter, M. B., and Bucker, H. P. (**1987**). "Gaussian beam tracing for computing ocean acoustic fields," J. Acoust. Soc. Am., **82**, 1349–1359. doi:10.1121/1.395269

Ramseur, J. L. (**2010**). "Deepwater Horizon oil spill: the fate of the oil," Cong Res Serv, **1**–**20**, R41531. Retrieved from https://fas.org/sgp/crs/misc/R41531.pdf

Rasmussen, M. H., Wahlberg, M., and Miller, L. A. (**2004**). "Estimated transmission beam pattern of clicks recorded from free-ranging white-beaked dolphins (Lagenorhynchus albirostris)," J. Acoust. Soc. Am., **116**, 1826–1831. doi:10.1121/1.1775274

Reeves, R. R., Lund, J. N., Smith, T. D., and Josephson, E. A. (**2011**). "Insights from whaling logbooks on whales, dolphins, and whaling in the Gulf of Mexico," Gulf Mex. Sci., **29**, 41–67. doi:10.18785/goms.2901.04

Seber, G. A. F. (**1982**). *The estimation of animal abundance and related parameters*, Charles W. Griffin, London, UK. doi:10.1002/iroh.19740590517

Townsend, C. H. (**1935**). "The distribution of certain whales as shown by the logbook records of American whaleships," Zool. Sci. Contrib. New York Zool. Soc., **19**, 3–50. Retrieved from https://www.biodiversitylibrary.org/part/203715

Ward, J. a., Thomas, L., Jarvis, S., Dimarzio, N., Moretti, D., Marques, T. a., Dunn, C., et al. (**2012**). "Passive acoustic density estimation of sperm whales in the Tongue of the Ocean, Bahamas," Mar. Mammal Sci., **28**, 444–455. doi:10.1111/j.1748-7692.2011.00560.x

Ward, J., Jarvis, S., Moretti, D., Morrissey, R., DiMarzio, N., Johnson, M., Tyack, P., et al. (**2011**). "Beaked whale (Mesoplodon densirostris) passive acoustic detection in increasing ambient noise," J. Acoust. Soc. Am., **129**, 662–669. doi:10.1121/1.3531844

Watwood, S. L., Miller, P. J. O., Johnson, M., Madsen, P. T., and Tyack, P. L. (**2006**). "Deep-diving foraging behaviour of sperm whales (Physeter macrocephalus)," J. Anim. Ecol., **75**, 814–825. doi:10.1111/j.1365-2656.2006.01101.x

Whitehead, H. (**2002**). "Estimates of the current global population size and historical trajectory for sperm whales," Mar. Ecol. Prog. Ser., **242**, 295–304. doi:10.3354/meps242295

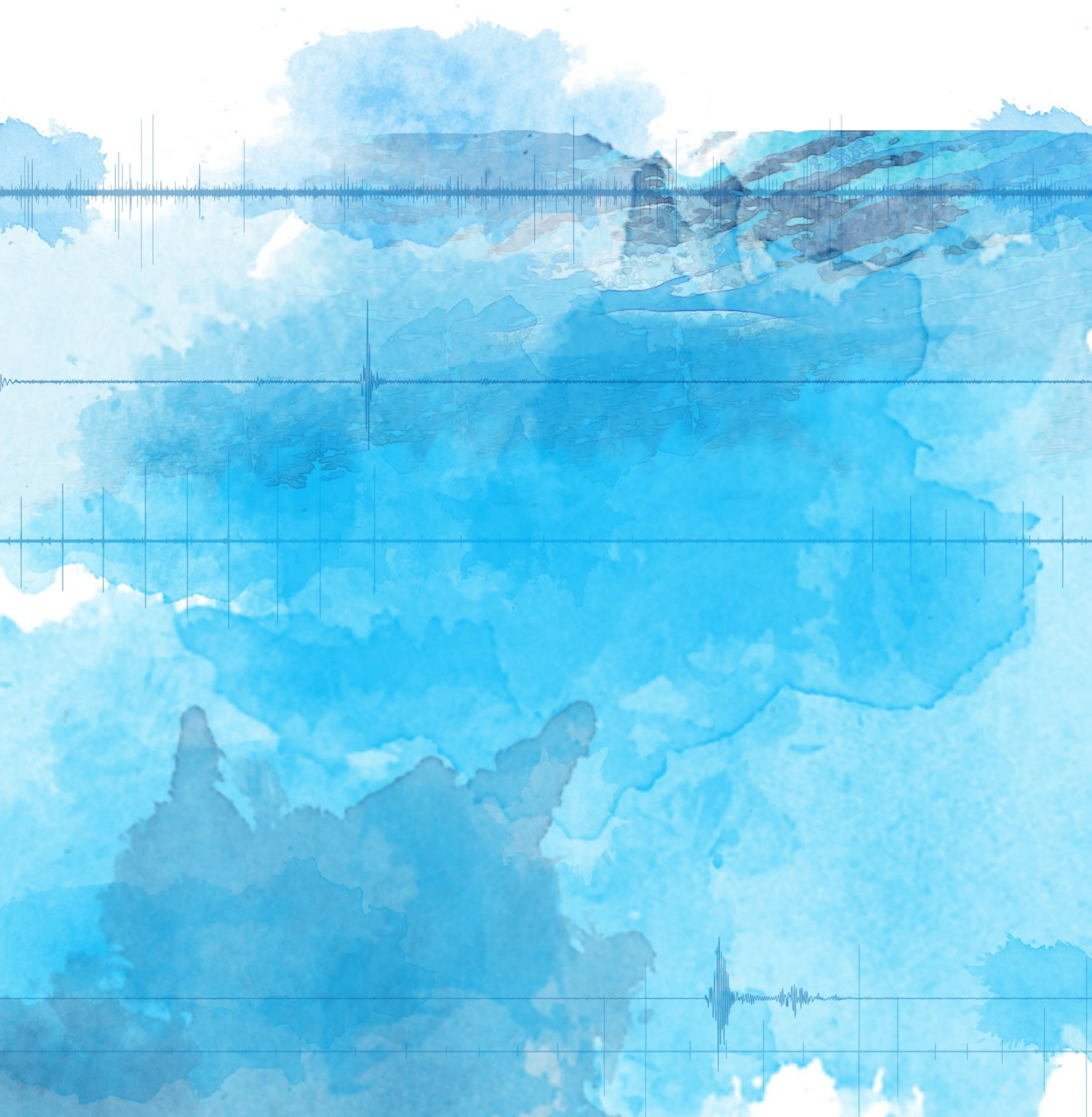
Wright, A. J., Soto, N. A., Baldwin, A. L., Bateson, M., Beale, C. M., and Clark, C. (**2007**). "Anthropogenic Noise as a Stressor in Animals: A Multidisciplinary Perspective," Int. J. Comp. Psychol., Retrieved from https://escholarship.org/uc/item/46m4q10x. Retrieved from https://escholarship.org/uc/item/46m4q10x

Zimmer, W. M. X., Johnson, M. P., Madsen, P. T., and Tyack, P. L. (**2005a**). "Echolocation clicks

of free-ranging Cuvier's beaked whales (Ziphius cavirostris)," J. Acoust. Soc. Am., **117**, 3919–3927. doi:10.1121/1.1910225

Zimmer, W. M. X., Tyack, P. L., Johnson, M. P., and Madsen, P. T. (**2005b**). "Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis," J. Acoust. Soc. Am., **117**, 1473. doi:10.1121/1.1828501

Conclusions and recommended future work



This thesis outlines advancements of passive acoustic monitoring methods as a framework to study deep-diving odontocetes. The framework provides efficient processing tools for long-term acoustic data and proves useful in quantifying population trends, fostering the applicability of these techniques for effective conservation strategies. The thesis presents substantial biological and ecological insights of the Gulf of Mexico's sperm whale population, which contributes to the overall understanding of how anthropogenic activities impact the population status and the dynamics of this endangered species and their habitat now and in the future. These advances have potential applications for solving many of the current challenges in long-term ecological monitoring.

The applicability of an automatic framework to classify echolocation clicks is presented within an evaluation framework (**Chapter 1**). The machine learning technique based on Gaussian Mixture Models generated suitable generalization abilities to perform automatic classification of Cuvier's beaked whale vocalizations from different regions and conditions. Difficulties in interpreting the results arose in part from the complexity of multiclass classification and the newly detected encounters that were not initially annotated on the ground truth data. Comparing model performance not only between regions, but also within regions, highlights the complexity in the generalization abilities of these models and the application of methodological approaches to conduct an evaluation of automatic techniques is recommended. This work provides the basis for implementing a framework with standardized evaluation which could advance the comparison and integration of new and improved techniques.

DetEdit, an open-source software program is presented (**Chapter 2**), which eases the analysis of acoustic data and minimizes the manual effort required to verify detections for species assessment from long-term monitoring studies. When implemented on the species-level analysis of odontocetes, an acceleration in the processing of acoustic data and removal of false positive detections was demonstrated. This open-source tool enables analysts to process acoustic data with minimal statistical computing experience and allows for the processing and characterization of aquatic and terrestrial stereotyped signals and characterization by spectral shape. It also facilitates the integration of advanced machine learning algorithms by building verified datasets and evaluating the results from automated methods. Although freely available, this tool is limited in the fact that it is implemented in a proprietary computing environment (*MATLAB*, Mathworks, Natick, MA). Further adaptations to freely available programming languages could increase the accessibility of this tool within the scientific community. In addition, implementation of an active online community would offer support and assistance with problems, further increasing user accessibility.

Several methods used to process the long-term passive acoustic data to study the sperm whale population in the Gulf of Mexico are provided throughout this thesis. A description of the method

used to detect sperm whale echolocation clicks, identify shipping passages, and verify encounters with the use of *DetEdit* is provided (**Chapter 2**).

A method to categorize sperm whale encounters into possible sex classes is described based on inter-click-interval distributions and the correlation between inter-pulse interval, acoustic body length, and inter-click interval (**Chapter 3**). Because the correlation was based on a manual analysis of a subset of data, uncertainty in estimates is likely. The size estimates for population structures could be improved using either automatic methods to sample stable inter-pulse intervals of all of the acoustic data or animal-borne acoustic recording tags to refine the relationship between size and acoustic behavior.

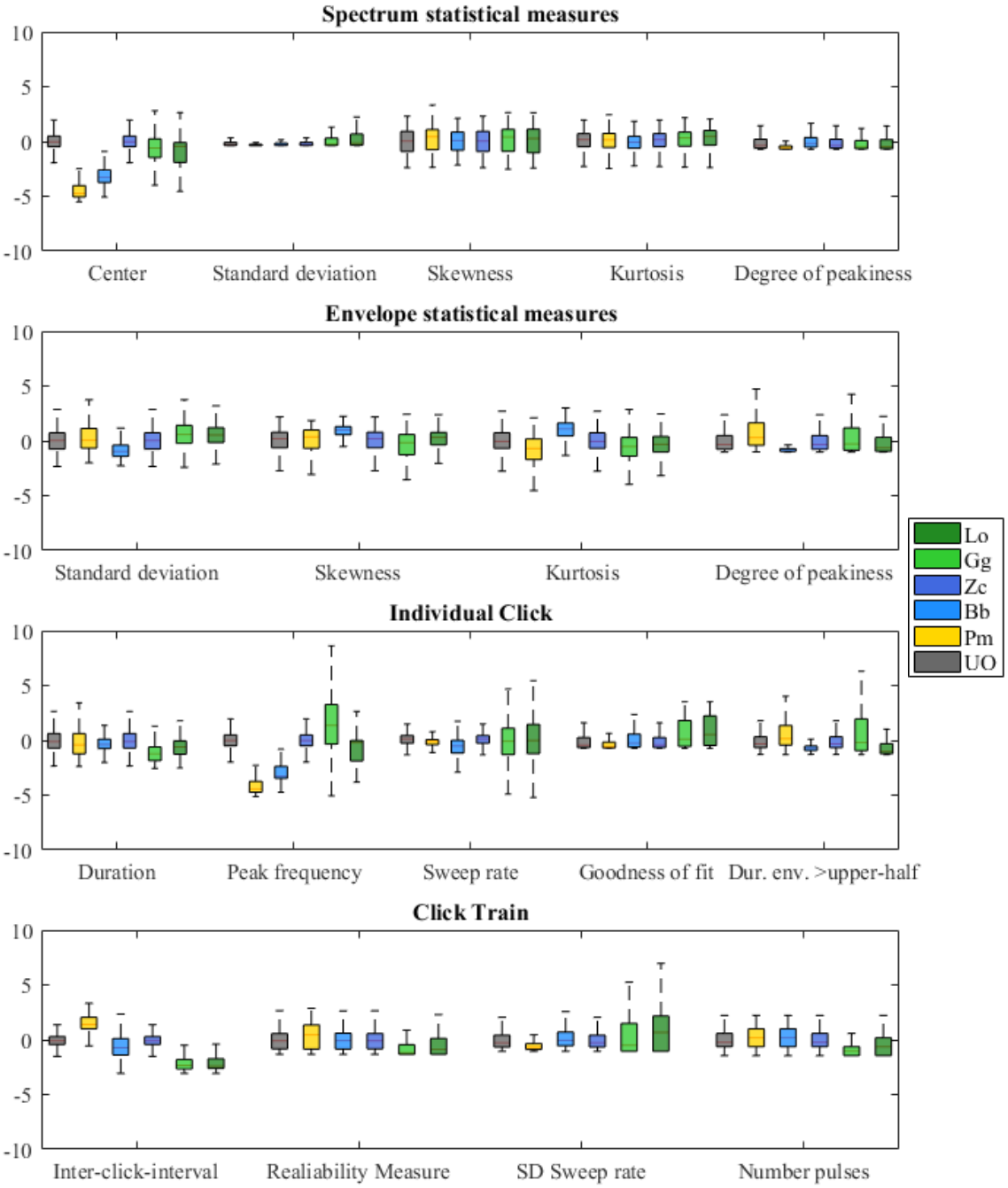
The analysis of the long-term acoustic data shed light on several aspects of the ecology of the Gulf of Mexico's sperm whale population. A greater understanding of sperm whale movement patterns, population structure, and population density trends, can provide a greater insight into the threats faced by this endangered species. Spatial and seasonal variation within the population structure of the Gulf of Mexico has been observed (**Chapter 3**), with high usage of the northern sites by adult females and their offspring. The results also reveal the presence of mature males in the area with an estimated size range consistent with adult males from the Atlantic region. This provides support to the hypothesis that males from the Atlantic move in and out of the Gulf of Mexico region, possibly to mate.

The challenges associated with using passive acoustics for density estimation, including the determination of population specific signal and behavioral parameters are illustrated based on the observed patterns of population structure (**Chapter 4**). Cue production rates accounting for population structure were approximated based on *in situ* data providing more accuracy, whereas signal and behavioral parameters were derived from acoustic tags and array recordings from previous studies and other populations yielding likely to biased densities. The evaluation of signal and behavioral parameter estimates per site should assess the influence of sex distribution on the estimation of the probability of detection by comparing model predictions and in situ localizations from an adequate array of sensors or by applying computer simulated sources. Although financially and logistically more challenging, tagging several animals (that are representative of the population structure) with acoustic devices would provide better estimates of the required parameters. Additionally, non-uniform distributions of the simulation framework should be utilized in an attempt to increase the accuracy of the detection function. Nonetheless, the population density estimates drawn from this work are in line with the reported densities of the global population but lower than the previously reported densities in the Gulf of Mexico. Several open questions remain regarding the drivers of the observed decline in sperm whale densities at the northern sites and contrary increases at the eastern site. Although this work expounds a correlation between the oil spill impact and commercial shipping with sperm whale population

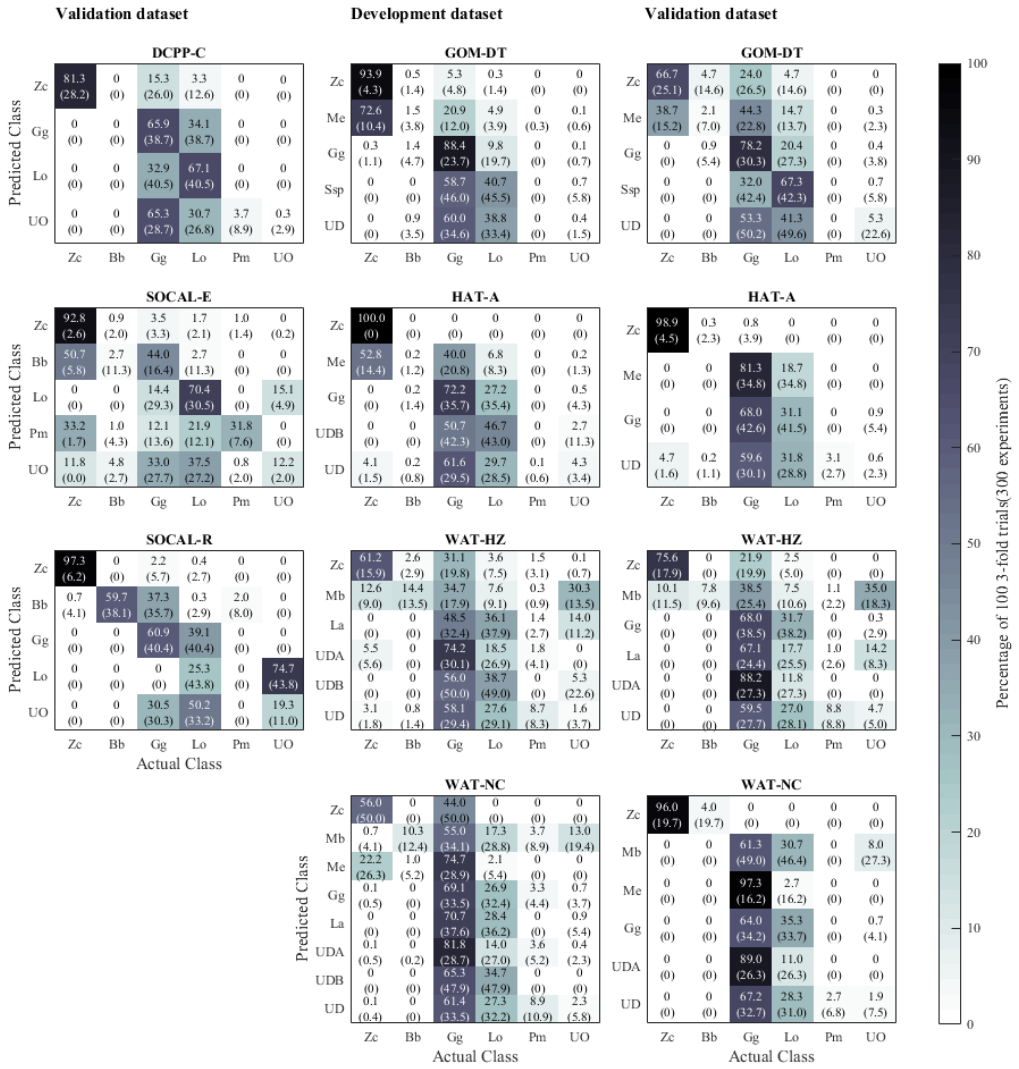
change, a better understanding of the sperm whale response to local conditions, due to both anthropogenic and natural processes is needed. The knowledge gathered through this work is part of a broader set of coordinated efforts to apply an unprecedented assessment of how the oil spill impacted a large and complex marine mammal community and their connected habitats. Additionally, this long-term time series could be integrated into predictive models for exploring relationships between environmental characteristics and density trends, activity, and behavior, in order to pinpoint environmental drivers of population change.

Monitoring marine mammals with passive acoustics provides cost-effective and sustainable methods to study ecosystem health, especially for areas where funding or accessibility is limited and long-term monitoring of marine mammals is necessary, like in the case of endangered species or in remote areas. Passive acoustic as an alternative method contributed to the data required for endangered and deficient-data species protection.

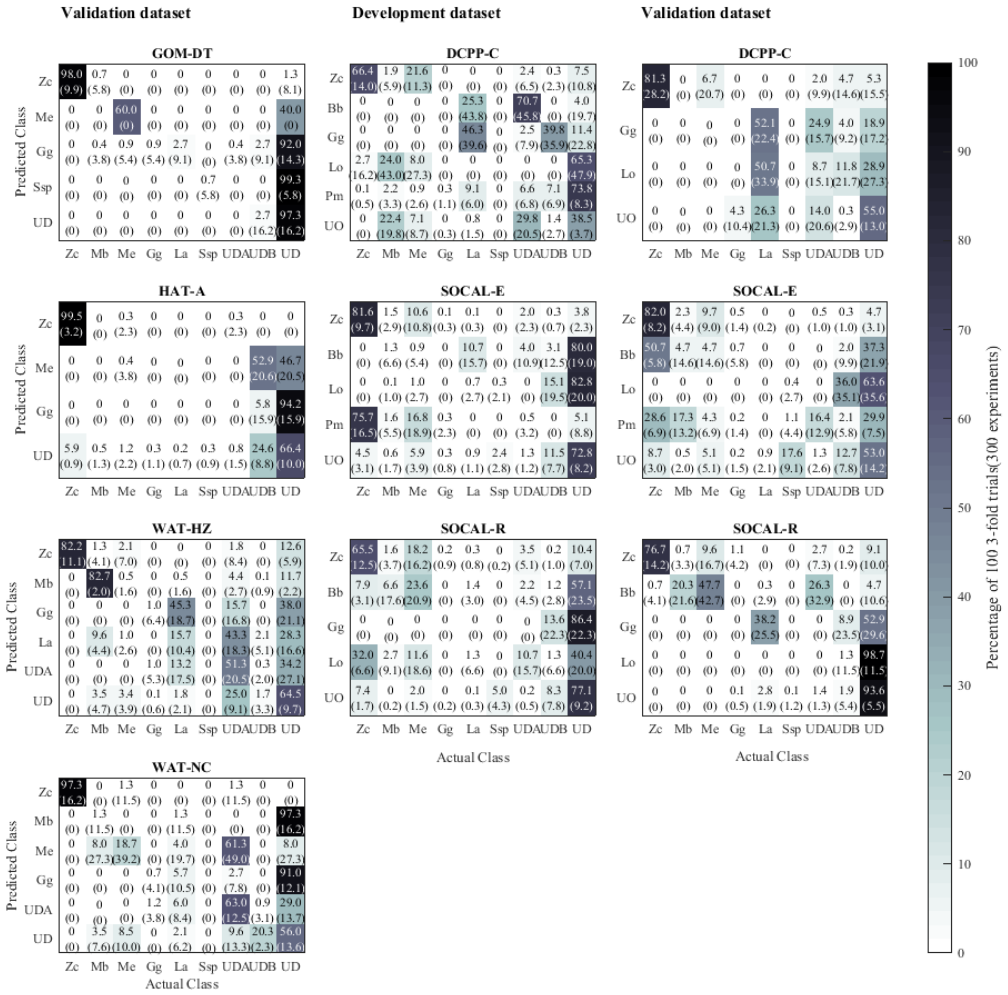
APPENDIX



Appendix I. Boxplot features per species. Each plot shows the features per feature group. Values are normalized to Cuvier's beaked whale class (*Zc*).



Appendix II. Mean confusion matrices (%) of Pacific GMM models' predictions of species classes from the Pacific validation dataset and both Atlantic datasets. Each confusion matrix shows the classification performance of each site. Elements of the matrix (numeric values and color-mapped visualization) represent the percentages of correctly (diagonal) or incorrectly (off-diagonal) classified species encounters by a GMM classifier among 75 experimental trials. Numbers in parenthesis are the standard deviation. Each column is normalized to a percentage, so values represented the percentages of 75 experiments for a given actual class that each possible class is predicted.



Appendix III. Mean confusion matrices (%) of Atlantic GMM models' predictions of species classes from the Pacific validation dataset and both Atlantic datasets. Each confusion matrix shows the classification performance of each site. Elements of the matrix (numeric values and color-mapped visualization) represent the percentages of correctly (diagonal) or incorrectly (off-diagonal) classified species encounters by a GMM classifier among 75 experimental trials. Numbers in parenthesis are the standard deviation. Each column is normalized to a percentage, so values represented the percentages of 75 experiments for a given actual class that each possible class is predicted.

Appendix IV. Parameters for rhythmic analysis algorithm detail as described in Zaugg et al. (2013).

Parameter name	Parameter value	Explanation
T_{ord}	20	Setting T_{ord} = 20 implies that selection of clicks contains up to 41 clicks. Hence 10 clicks from the target train would still be included in section even if they were interleaved with 31 nuisance clicks.
$[\tau_{low}, \tau_{up}]$	[0.01, 1] s	The fastest expected beaked whale click trains have an ICI typically around 0.2s and dolphin and beaked whale lowest expected click train have an ICI of 0.9s. (Frantzis et al., 2002; Zimmer et al., 2005). Hence 1s is a conservative upper bound.
$[T_{low}, T_{up}]$	[0.01, 5] s	T_{low} chosen for the same reason as τ_{low}. The value of T_{up} allows for 11 slow beaked whale clicks (ICI = 0.9s) to be included in the following selection of clicks that fit with the time delays of a typical ICI click train from the focal pulse.
N	1.5ms	The duration of beaked whale clicks and dolphin clicks is below 1ms. A window of 1.5ms covers this duration.
$[F_{low}, F_{up}]$	[5, 80] kHz	Beaked whale clicks and dolphin clicks have the most energy below 80 kHz. The band below 5kHz contains mostly noise and sperm whales.
b	2	Beaked whale clicks and dolphin clicks typically cover ~80 kHz bandwidth. A block size of 2 bins, which corresponds to ~1.3 kHz, seems adequate to catch the specific shape of the spectra.
N_{del}	10	Hence, final selection of clicks contains up to 10 clicks.
Q	0.1	Based on the assumption that ICI jitter of beaked whale clicks trains is generally smaller than 20%.
R_{tol}	0.2	Based on the same assumption as above that ICI jitter of beaked whale click trains is generally with in 20%.
U_{diss}		Defined by the dissimilarity metric of the selected neighbor clicks. Threshold is based on the median - median absolute deviation (Leys et al., 2013).

Appendix V. Evaluation measures from development datasets. Values are mean percentages and standard deviation of 100 3-fold experimental trials.

Species	Location	Precision	Recall	F-score	Truth Coverage < 10min	Truth Coverage ≥ 10min
Pacific Model						
Zc	DCPP-C	93.6 (8.9)	98.4 (5.8)	95.6 (5.8)	71.9 (31.9)	88.5 (0.0)
	SOCAL-E	100.0 (0.0)	96.2 (1.8)	98.1 (1.0)	71.5 (31.2)	94.3 (8.7)
	SOCAL-R	100.0 (0.2)	88.5 (3.7)	93.9 (2.1)	66.1 (32.4)	93.0 (11.2)
Atlantic Model						
Zc	GOM-DT	100.0 (0.2)	94.9 (5.6)	97.3 (3.4)	70.8 (33.8)	92.0 (17.7)
	HAT-A	98.1 (2.6)	97.1 (5.6)	97.5 (3.3)	85.9 (9.5)	95.6 (3.1)
	WAT-HZ	99.9 (0.6)	79.4 (10.3)	88.1 (7.2)	59.8 (35.4)	81.6 (20.0)
	WAT-NC	100.0 (0.0)	68.7 (46.5)	68.7 (46.5)	65.4 (0.0)	
Mb	WAT-HZ	99.3 (2.4)	94.5 (3.0)	96.8 (1.9)		89.6 (17.2)
	WAT-NC	80.3 (27.5)	100.0 (0.0)	86.0 (20.1)	100.0 (0.0)	
Me	GOM-DT	98.4 (3.5)	97.1 (1.4)	97.7 (2.0)	71.4 (34.6)	96.8 (2.5)
	HAT-A	97.1 (4.4)	97.0 (3.4)	96.9 (2.6)	78.9 (25.9)	95.7 (7.7)
	WAT-NC	82.7 (15.9)	95.5 (9.4)	87.2 (9.9)	93.7 (9.6)	91.1 (6.0)

Zc: *Ziphius cavirostris*; **Mb**: *Mesoplodon bidens*; **Me**: *M. europaeus*

Appendix VI. Evaluation measure of Cuvier's beaked whale (*Zc*) models for the validation datasets. Values are mean percentages and standard deviation of 75 experimental trials.

Location	Pacific Model				Atlantic Model			
	Precision	Recall	F-score	Truth Coverage	Precision	Recall	F-score	Truth Coverage
Pacific dataset								
DCPP-C	94.4 (13.0)	81.3 (28.2)	83.1 (22.9)	79.9 (38.4)	100.0 (0.0)	81.3 (28.2)	86.2 (23.3)	79.9 (38.4)
SOCAL-E	77.1 (1.8)	90.1 (2.1)	83.1 (1.1)	78.8 (34.5)	79.4 (2.8)	80.1 (7.5)	79.5 (4.0)	70.3 (40.9)
SOCAL-R	86.8 (5.0)	97.3 (6.2)	91.5 (3.0)	89.9 (18.0)	98.2 (5.18)	76.7 (14.2)	85.1 (11.0)	70.5 (40.1)
Atlantic dataset								
GOM-DT	20.7 (6.2)	77.3 (18.3)	32.8 (7.9)	70.5 (38.6)	36.6 (13.7)	98.7 (6.6)	52.1 (12.4)	91.8 (12.4)
HAT-A	36.1 (4.4)	98.9 (4.5)	52.7 (4.7)	83.4 (24.9)	33.9 (3.7)	99.5 (3.2)	50.4 (3.7)	84.1 (23.9)
WAT-HZ	71.6 (14.5)	68.0 (16.1)	68.1 (13.0)	49.1 (40.6)	84.7 (5.5)	74.0 (10.0)	78.6 (7.3)	54.6 (39.8)
WAT-NC	100.0 (0.0)	93.0 (19.7)	96.0 (19.7)	90.9 (18.7)	100.0 (0.0)	97.3 (16.2)	97.3 (16.2)	92.1 (15.4)

Appendix VII. Evaluation measures for Cuvier's beaked whale (*Zc*) models for development data of opposite region. Values are mean percentages and standard deviation of 75 experimental trials.

Location	Precision	Recall	F-score	Truth Coverage
Atlantic Model				
DCPP-C	99.6 (3.2)	76.7 (22.0)	84.5 (17.4)	67.4 (42.4)
SOCAL-E	81.3 (3.9)	80.3 (8.2)	80.4 (4.4)	70.6 (40.0)
SOCAL-R	79.5 (19.4)	69.1 (15.1)	72.7 (15.4)	54.1 (44.9)
Pacific Model				
GOM-DT	26.9 (8.2)	84.4 (15.0)	40.7 (10.1)	81.0 (32.6)
HAT-A	33.3 (5.4)	99.5 (3.2)	49.6 (6.1)	90.5 (13.5)
WAT-HZ	61.0 (15.9)	60.1 (16.7)	60.0 (14.0)	43.4 (41.5)
WAT-NC	69.7 (41.5)	76.0 (42.9)	71.7 (38.7)	72.1 (40.7)

There's music in the deep:
It is not in the surf's rough roar,
Nor in the whispering, shelly shore
They are but earthly sounds, that tell
How little of the sea nymph's shell,
That sends its loud, clear note abroad,
Or winds its softness through the flood,
Echoes through groves with coral gay,
And dies, on spongy banks, away.
There's music in the deep.

The Deep, John G. C. Brainard